

Secret Communication over Broadcast Erasure Channels with State-feedback

László Czap, Vinod M. Prabhakaran, Christina Fragouli, Suhas Diggavi

Abstract

We consider a 1-to- K communication scenario, where a source transmits private messages to K receivers through a broadcast erasure channel, and the receivers feed back strictly causally and publicly their channel states after each transmission. We explore the achievable rate region when we require that the message to each receiver remains secret - in the information theoretical sense - from all the other receivers. We characterize the capacity of secure communication in all the cases where the capacity of the 1-to- K communication scenario without the requirement of security is known. As a special case, we characterize the secret-message capacity of a single receiver point-to-point erasure channel with public state-feedback in the presence of a passive eavesdropper.

We find that in all cases where we have an exact characterization, we can achieve the capacity by using linear complexity two-phase schemes: in the first phase we create appropriate secret keys, and in the second phase we use them to encrypt each message. We find that the amount of key we need is smaller than the size of the message, and equal to the amount of encrypted message the potential eavesdroppers jointly collect. Moreover, we prove that a dishonest receiver that provides deceptive feedback cannot diminish the rate experienced by the honest receivers.

We also develop a converse proof which reflects the two-phase structure of our achievability scheme. As a side result, our technique leads to a new outer bound proof for the non-secure communication problem.

I. INTRODUCTION

Wireless communication channels are easier to eavesdrop and harder to secure – even towards unintentional eavesdroppers. As an example, consider a sender, Alice, who wants to send private messages to multiple (say three) receivers, Bob, Calvin and David, within her transmission radius, and assume public feedback from the receivers to Alice. When Alice broadcasts a message W_1 intended for Bob, Calvin and David should also try to overhear, as the side information they possibly collect can enable Alice to make her following broadcast transmissions more efficient; but then, this collected side information would allow Calvin and David to learn parts of Bob's message. Even worse, Calvin and David could try to put together the parts they overheard, to extract increased information about Bob's message. Can we, in such a setting, keep the message for each user information theoretically secure from the other users, even if these nodes collaborate? Moreover, can we do so, when the users can only communicate through shared wireless broadcast channels?

In this paper, we answer these questions when communication happens through a broadcast packet erasure channel with public feedback. In particular, we assume that the receivers acknowledge through a public channel whether or not they correctly received packets; this is a natural assumption that is aligned with the operation of current wireless standards. Recent results justify the relevance of such an erasure model, e.g. [4] shows that a state dependent Gaussian channel can be viewed as a packet erasure channel. We exactly characterize the capacity region in all the cases where the problem has been solved with no security constraints, namely, the 2-user, 3-user, symmetric K -user, and one-sidedly fair K -user [5], [6] cases. For each such case, we present a new outer bound and a polynomial time achievability scheme that matches it.

Our achievability schemes operate in two phases: in the first phase we efficiently generate secure keys between the source and each of the receivers, while in the second we judiciously use these keys for encryption. In both phases, we exploit channel properties to make our protocols efficient in terms of achieved rates.

In the first phase, we make use of a fundamental observation by Maurer [7]: different receivers have different looks on the transmitted signals, and we can build on these differences with the help of feedback to create secret keys [8], [9]. For example, if the sender – call her Alice – transmits random packets through independent erasure channels with erasure probability 0.5, there would be a good fraction of them (approximately 25%) that only one of two users receives, and we can transform this common randomness between Alice and the given user to a key using privacy amplification [7]–[10]. Testbed implementations have demonstrated that a secret-key rate of several tens of Kbps is achievable by exploiting erasures in a practical wireless setting [11], [12].

In the second phase, we use the generated keys to transmit private messages. A naive approach is to generate secret keys *of the same size* as the size of the respective private messages, and then use these keys as one-time pads. However, this is too pessimistic in our case: the other users are going to receive only a fraction of the encrypted messages. Thus, *we can use keys*

L. Czap is with EPFL, Switzerland. Email: laszlo.czap@epfl.ch

V. M. Prabhakaran is with TIFR, India. Email: vinodmp@tifr.res.in

C. Fragouli is with UCLA, USA and EPFL, Switzerland. Email: christina.fragouli@epfl.ch

S. Diggavi is with UCLA, USA. Email: suhas@ee.ucla.edu

This paper was presented in parts at the IEEE Information Theory Workshop (ITW 2011) [1], at the IEEE International Symposium on Information Theory (ISIT 2012) [2] and at the IEEE International Symposium on Network Coding (NetCod 2013) [3].

of smaller size than the messages, and still be secure. To build on this observation, feedback is useful; knowing which packets a given user has successfully received, allows us to decide what to transmit next, so that we preserve secrecy from the others.

Our schemes assume that the users provide honest feedback, but they can be extended when this is not the case. To illustrate, for the special case of $K = 2$, we design a scheme that provides secrecy for an honest user even if the other user provides potentially false acknowledgments. Interestingly, we find that the same rate is achievable against dishonest users as against honest-but-curious users. We note however that, although our scheme against dishonest users is optimal in terms of achieved rates, its security relies on the uniform distribution of the message for the dishonest party. From a practical perspective such an assumption is potentially too restrictive, which motivates us to define the new notion of distribution independent security and design a scheme that fulfills this latter, stronger security notion. We also take the opportunity to investigate the relation between different notions of security. In particular, following [13] we show equivalence between our security notions and semantic security.

To prove the optimality of our achievability schemes, we derive a new impossibility result for the secure 1-to- K message transmission problem, that applies for all values of K and any channel parameters. Our converse proof introduces a new technique that explicitly utilizes a balance between generated and consumed keys, indicating that generating and using keys is a natural strategy. As a side result, we provide a new proof for the known outer bound of the non-secure rate region derived in [5], [6].

Finally, our work also provides the secure-message capacity of a point-to-point erasure channel with public state-feedback in the presence of a passive eavesdropper Eve. When no feedback is available, secrecy is achievable only if the legitimate receiver has a channel of larger capacity than the wiretapper [14]. If public feedback is available, the work of Maurer proves that a nonzero secret key generation rate is achievable as long as the eavesdropper does not have an error free channel [7]; in this work we show we can also securely send specific messages, yet at rates lower than it is possible for key-generation. We illustrate this in Figure 3.

To the best of our knowledge, this work provides the first characterization of secret-message capacity with limited public feedback for a non-trivial setup.

II. RELATED WORK

We distinguish between secret-key exchange, where Alice and Bob wish to agree on a common secret securely from a passive eavesdropper, Eve, and secret-message exchange, where Alice wants to send a specific message to Bob securely from Eve.

For the wiretap channel, when there is no feedback from Bob, the rates for secret key and secret-message exchange coincide; in his seminal work, Wyner derived the achievable rates of secure communication over a noisy point-to-point channel and showed that unless Eve has a worse channel than Bob, the secure communication rate is zero [14]. These results on the wiretap channel were generalized in several directions, see e.g. [15], [16]. The works of Maurer, Ahlswede and Csiszár have shown that in contrast, if public feedback is available, we can achieve non-zero key generation rates even if Eve has a better channel than Bob, thus establishing that public feedback can significantly increase the achievable secure key generation rates [7], [17]–[19]. The wiretap channel with secure (inaccessible to Eve) feedback has been studied in [20]–[22]. Our work focuses on the message exchange problem, and shows that for the case of erasure channels, public feedback also increases the secure message exchange rate; yet, in this case, the message exchange capacity is smaller than the secret-key exchange capacity. Our results go beyond the point-to-point channel to the case of sending private messages to multiple receivers. In addition, in [7], [19] a public channel of infinite capacity is assumed, whereas we restrict the public feedback to a short channel-state acknowledgment, which is more practical.

Recently there has been a number of interesting works that build on physical channel characteristics (such as channel reciprocity) to derive key generation schemes [23], [24]. Our work focuses on erasure channels and on the message sending problem [1], [2]. Secret-message exchange has also been studied over networks in the case where there are no channel errors [25]; in our setup, the channel variability is what makes secrecy possible.

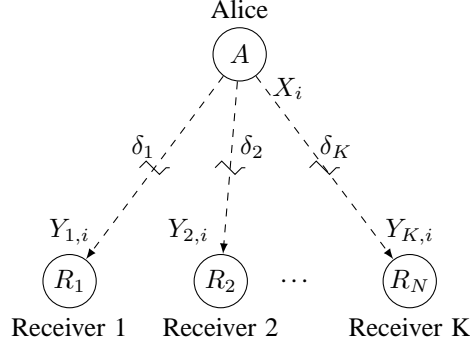
Finally, as mentioned in the introduction, the use of feedback and broadcast for private message transmission without security requirements has been studied in [6], [26], [27].

III. DEFINITIONS AND BACKGROUND

A sender, Alice, wants to send private messages to a set of K receivers: she wants to send message W_j to receiver j , so that, no other receiver learns W_j , even if all other receivers collude.

A. Communication model

Communication takes place over a 1-to- K broadcast erasure channel, with input at Alice and an output at each of the K receivers. We illustrate the setting in Figure 1. We also summarize our frequently used notations in Table I. The channel input alphabet consists of all possible vectors of length L over a finite field \mathbb{F}_q . For convenience, we usually call such a vector a

Figure 1. 1-to- K broadcast erasure channelTable I
SUMMARY OF NOTATION

$X_i, Y_{j,i}$	The i th input and outputs of the channel
S_i, S_i^*	The actual and the acknowledged state of the channel in the i th transmission
δ_j	Erasure probability for receiver j
W_j	Private message for receiver j
P_{W_i}, P_{W_i, W_j}	Distribution and joint distribution of messages W_i, W_j
σ_1, σ_2	Acknowledging strategy of a dishonest user
N_j	Size of W_j expressed in number of packets
R_j	Secret message rate for receiver j
Θ_A, Θ_j	Private randomness of Alice and of receiver j

packet. Throughout the paper we express entropy and rate in terms of packets. This enables us to omit the constant factor $L \log q$.

The broadcast channel is made up of K independent¹ component erasure channels with packet erasure probabilities $\delta_1, \delta_2, \dots, \delta_K$. Below we define the channel formally (see Eq. (1)-(2)). We assume that the receivers send public acknowledgments after each transmission stating whether or not they received the transmission correctly. By *public* we mean that the acknowledgments are available not only for Alice but for all other receivers as well. We assume that some authentication method prevents the receivers from forging each other's acknowledgments. Also, receivers learn each other's acknowledgments causally, after they have revealed their own.

Let $S_i \in 2^{\{1,2,\dots,K\}}$ denote the state of the channel in the i th transmission. S_i collects the indices of receivers with correct reception. In the first place, we assume that acknowledgments are honest. We relax this assumption for the special case of $K = 2$.

We denote by X_i the i th transmission over the channel, and by $Y_i = (Y_{1,i}, Y_{2,i}, \dots, Y_{K,i})$ the corresponding outputs observed by the receivers. We use X^n to denote the vector (X_1, X_2, \dots, X_n) . We use a similar notation for other vectors as well. Formally, the channel behavior is defined as

$$\Pr\{Y_i|X_i\} = \prod_{j=1}^K \Pr\{Y_{j,i}|X_i\} \quad (1)$$

$$\forall j \in \{1, 2, \dots, K\} : \Pr\{Y_{j,i}|X_i\} = \begin{cases} 1 - \delta_j, & Y_{j,i} = X_i \\ \delta_j, & Y_{j,i} = \perp, \end{cases} \quad (2)$$

where \perp is the symbol of erasure.

We assume that all participants can generate private randomness at unlimited rate. We denote the private randomness of the sender Θ_A , while Θ_j is the private randomness of user j . Variables Θ_A and every Θ_j are independent from all other randomness in the system.

B. Reliability and security – honest-but-curious users

An $(n, \epsilon, N_1, N_2, \dots, N_K)$ scheme sends message W_j which consist of N_j packets of length L to receiver j using n transmissions from Alice with error probability smaller than ϵ . We denote $W = (W_1, W_2, \dots, W_K)$ the set of all messages.

¹We assume independence for simplicity, but as long as the statistical behavior of the channel is known, our results can be easily generalized.

Definition 1. An $(n, \epsilon, N_1, N_2, \dots, N_K)$ scheme for the 1-to- K broadcast channel consists of the following components: (a) message alphabets $\mathcal{W}_j = \mathbb{F}_q^{LN_j}$, $j = 1, 2, \dots, K$, (b) encoding maps $f_i(\cdot)$, $i = 1, 2, \dots, n$, and (c) decoding maps $\phi_j(\cdot)$, $j = 1, 2, \dots, K$, such that the inputs to the channel are

$$X_i = f_i(W, \Theta_A, S^{i-1}), \quad i = 1, 2, \dots, n, \quad (3)$$

where messages $W_j \in \mathcal{W}_j$ are arbitrary messages in their respective alphabets and Θ_A is the private randomness Alice has access to. Further, it provides decodability for each receiver, that is

$$\forall 1 \leq j \leq K : \Pr\{\phi_j(Y_j^n S^n) \neq W_j\} < \epsilon \quad (4)$$

is satisfied.

Definition 2. The capacity region of the 1-to- K broadcast erasure channel $\mathcal{R}^K \subset \mathbb{R}_+^K$ is the set of rate tuples for which for every $\epsilon > 0$ there exists an $(n, \epsilon, N_1, N_2, \dots, N_K)$ scheme that satisfies

$$\forall 1 \leq j \leq K : R_j - \epsilon < \frac{1}{n} N_j \quad (5)$$

for all $1 \leq j \leq K$.

The following definition extends Definition 1 with a security requirement.

Definition 3. An $(n, \epsilon, N_1, N_2, \dots, N_K)$ scheme is secure against honest-but-curious users if in addition to (3)-(4) the following also holds for all $1 \leq j \leq K$:

$$\max_{P_W} I(W_j; Y_{-j}^n S^n \Theta_{-j}) < \epsilon, \quad (6)$$

where the maximum is taken over all possible joint message distributions and Y_{-j}^n is a shorthand for $Y_1^n, \dots, Y_{j-1}^n, Y_{j+1}^n, \dots, Y_K^n$. Similarly, Θ_{-j} is $\Theta_1, \dots, \Theta_{j-1}, \Theta_{j+1}, \dots, \Theta_K$.

Definition 4. The secret-message capacity region of the 1-to- K broadcast erasure channel $\mathcal{R}_H^K \subset \mathbb{R}_+^K$ is the set of rate tuples for which for every $\epsilon > 0$ there exists an $(n, \epsilon, N_1, N_2, \dots, N_K)$ scheme that is secure against honest-but-curious users and satisfies

$$\forall 1 \leq j \leq K : R_j - \epsilon < \frac{1}{n} N_j. \quad (7)$$

Following [5] we distinguish two special cases.

Definition 5. We call the channel symmetric if the erasure probabilities are all the same: $\delta_i = \delta_j, \forall 1 \leq i, j \leq K$.

Definition 6. We call a rate vector one-sidedly fair if $\delta_i \geq \delta_j$ for $i \neq j$ implies

$$R_i \delta_i \geq R_j \delta_j. \quad (8)$$

C. Dishonest users

For the special case of $K = 2$ we introduce further, stronger notions of security. In particular, we aim to provide secrecy against receivers who might acknowledge dishonestly. For convenience, we call the receivers Bob and Calvin. A *dishonest user* can produce dishonest acknowledgments as a (potentially randomized) function of all the information he has access to when producing each acknowledgment (this includes all the packets and the pattern of erasures he received up to and including the current packet he is acknowledging and the acknowledgments sent by the other user over the public channel up to the previous packet). In the following \mathcal{S}_1 and \mathcal{S}_2 denote the set of all possible acknowledging strategies of Bob and Calvin respectively and $\sigma_1 \in \mathcal{S}_1$ and $\sigma_2 \in \mathcal{S}_2$ denote their elements.

We do not provide any guarantees for a dishonest user, hence at most one of the two receivers can be dishonest, otherwise the problem would not be meaningful. Of course, the sender, Alice is not aware of which user is dishonest, otherwise she could simply ignore the dishonest party.

Our security definition for the case of honest-but-curious users does not depend on the joint distribution of the messages. In contrast, we define security against a dishonest user under three different assumptions on the joint message distribution that correspond to different levels of security. First, we assume that messages are independent and uniformly distributed (see Definitions 8-9). This models the case when messages are properly source-coded and the users have no control on them. Second, we relax any assumption on the message distribution (see Definitions 10-11). This model even allows that the dishonest user selects arbitrarily the joint distribution of the messages. Third, we assume that messages are independent and the message of the dishonest user is uniformly distributed, but we do not make any assumption on the message distribution of the honest user (see Definitions 12-13). In this model, the dishonest user might choose the distribution of only the other user's message.

According to another interpretation, in this case, the dishonest user might have side information about the message distribution of the honest user. Security in this last model also ensures resistance against a chosen-plaintext attack.

We denote by S_i^* the i th channel state based on the acknowledgments from Bob and Calvin. If there is a dishonest user, then potentially $S_i \neq S_i^*$, thus Alice and the honest party do not have access to the true channel states. We need to modify Definition 1 for $K = 2$ accordingly:

Definition 7. An (n, ϵ, N_1, N_2) scheme for the two user message transmission problem consists of the following components: (a) message alphabets $\mathcal{W}_1 = \mathbb{F}_q^{LN_1}$ and $\mathcal{W}_2 = \mathbb{F}_q^{LN_2}$, (b) encoding maps $f_i(\cdot)$, $i = 1, 2, \dots, n$, and (c) decoding maps $\phi_1(\cdot)$ and $\phi_2(\cdot)$, such that if the inputs to the channel are

$$X_i = f_i(W_1, W_2, \Theta_A, S^{*i-1}), \quad i = 1, 2, \dots, n, \quad (9)$$

where $W_1 \in \mathcal{W}_1$ and $W_2 \in \mathcal{W}_2$ are arbitrary messages in their respective alphabets and Θ_A is the private randomness Alice has access to. Then, provided the receivers acknowledge honestly,

$$\Pr\{\phi_1(Y_1^n S^{*n}) \neq W_1\} < \epsilon, \text{ and} \quad (10)$$

$$\Pr\{\phi_2(Y_2^n S^{*n}) \neq W_2\} < \epsilon. \quad (11)$$

Security under uniform message distribution: The following definition extends Definition 7 with a security requirement assuming that messages are independent and uniformly distributed.

Definition 8. An (n, ϵ, N_1, N_2) scheme is said to be secure against a dishonest user under uniform message distribution, if it guarantees decodability and security for an honest user even if the other user is dishonest (as defined above). That is, if W_1 and W_2 are independent and both are uniformly distributed, then when Bob is honest,

$$\max_{\sigma_2} \Pr\{\phi_1(Y_1^n S^{*n}) \neq W_1\} < \epsilon \quad (12)$$

$$\max_{\sigma_2} I(W_1; Y_2^n S^n \Theta_2) < \epsilon \quad (13)$$

are satisfied, and when Calvin is honest,

$$\max_{\sigma_1} \Pr\{\phi_2(Y_2^n S^{*n}) \neq W_2\} < \epsilon \quad (14)$$

$$\max_{\sigma_1} I(W_2; Y_1^n S^n \Theta_1) < \epsilon. \quad (15)$$

are satisfied. The maxima are taken over all adversarial acknowledging strategies.

Definition 9. The rate region $\mathcal{R}_{uDH}^2 \subset \mathbb{R}_+^2$ is the set of rate pairs for which for every $\epsilon > 0$ there exists an (n, ϵ, N_1, N_2) scheme that is secure against a dishonest user under uniform message distribution and satisfies

$$R_1 - \epsilon < \frac{1}{n} N_1 \text{ and } R_2 - \epsilon < \frac{1}{n} N_2. \quad (16)$$

Distribution independent security: Below, we relax the assumption on the message distribution. This leads to a stronger notion of security that we call distribution independent security.

Definition 10. An (n, ϵ, N_1, N_2) scheme is said to provide distribution independent security, if it guarantees decodability and security for the honestly acknowledging user (or users) independently of the joint distribution P_{W_1, W_2} of (W_1, W_2) . That is, if Bob is honest,

$$\max_{P_{W_1, W_2}, \sigma_2} \Pr\{\phi_1(Y_1^n S^{*n}) \neq W_1\} < \epsilon \quad (17)$$

$$\max_{P_{W_1, W_2}, \sigma_2} I(W_1; Y_2^n S^n \Theta_2 | W_2) < \epsilon \quad (18)$$

are satisfied, and if Calvin is honest, then

$$\max_{P_{W_1, W_2}, \sigma_1} \Pr\{\phi_2(Y_2^n S^{*n}) \neq W_2\} < \epsilon \quad (19)$$

$$\max_{P_{W_1, W_2}, \sigma_1} I(W_2; Y_1^n S^n \Theta_1 | W_1) < \epsilon. \quad (20)$$

are satisfied.

Definition 11. The rate region $\mathcal{R}_{DIS}^2 \subset \mathbb{R}_+^2$ is the set of rate pairs for which for every $\epsilon > 0$ there exists an (n, ϵ, N_1, N_2) scheme that provides distribution independent security and satisfies

$$R_1 - \epsilon < \frac{1}{n} N_1 \text{ and } R_2 - \epsilon < \frac{1}{n} N_2. \quad (21)$$

Security against a user with side information: In the following two definitions we assume that a dishonest user can choose the message distribution of the other user, but not his own.

Definition 12. An (n, ϵ, N_1, N_2) scheme is said to be secure against a dishonest user with side information, if it guarantees decodability and security for an honest user even if the other user is dishonest (as defined above) and can choose the message distribution of the honest user. That is, if W_1 and W_2 are independent, then if W_2 is uniformly distributed and Bob is honest,

$$\max_{P_{W_1, \sigma_2}} \Pr\{\phi_1(Y_1^n S^{*n}) \neq W_1\} < \epsilon \quad (22)$$

$$\max_{P_{W_1, \sigma_2}} I(W_1; Y_2^n S^n \Theta_2) < \epsilon \quad (23)$$

are satisfied, whereas if W_1 is uniformly distributed and Calvin is honest,

$$\max_{P_{W_2, \sigma_1}} \Pr\{\phi_2(Y_2^n S^{*n}) \neq W_2\} < \epsilon \quad (24)$$

$$\max_{P_{W_2, \sigma_1}} I(W_2; Y_1^n S^n \Theta_1) < \epsilon \quad (25)$$

are satisfied. The maxima are taken over all adversarial acknowledging strategies and all possible message distributions of the honest user.

Definition 13. The rate region $\mathcal{R}_{DH}^2 \subset \mathbb{R}_+^2$ is the set of rate pairs for which for every $\epsilon > 0$ there exists an (n, ϵ, N_1, N_2) scheme that is secure against a dishonest user with side information and satisfies

$$R_1 - \epsilon < \frac{1}{n} N_1 \text{ and } R_2 - \epsilon < \frac{1}{n} N_2. \quad (26)$$

D. Non-secure 1-to- K broadcast

Before summarizing our results, we restate the result from [5], [6] that characterizes \mathcal{R}^K in the known cases. Let π denote a permutation of $\{1, 2, \dots, K\}$ and π_i the i th element of the permutation.

Theorem 1. For $K \leq 3$ or for a symmetric channel with $K > 3$ or for a one-sidedly fair rate tuple $(R_1, \dots, R_K) \in \mathbb{R}_+^K$ with $K > 3$, the capacity region \mathcal{R}^K of the 1-to- K broadcast erasure channel with state-feedback is characterized by the following inequality:

$$\max_{\pi} \sum_{i=1}^K \frac{R_{\pi_i}}{1 - \prod_{k=1}^i \delta_{\pi_k}} \leq 1, \quad (27)$$

where the maximization is taken over all permutations π of $\{1, \dots, K\}$.

Further, it was shown in [5] and [6] that (27) is an outer-bound for \mathcal{R}^K in all cases.

Theorem 2. Any rate tuple $(R_1, \dots, R_K) \in \mathbb{R}_+^K$ in \mathcal{R}^K satisfies (27).

IV. SUMMARY OF RESULTS

In this section we provide an overview of the results that we present in this paper.

A. Honest-but-curious users

Our main result for honest-but-curious users is the characterization of the secret-message capacity region \mathcal{R}_H^K for sending private messages to K receivers over a broadcast erasure channel, for all the cases where the capacity region without secrecy constraints \mathcal{R}^K has been characterized, namely, the 2-user, 3-user, symmetric K -user and one-sidedly fair K -user cases.

For all the mentioned cases, when the capacity region \mathcal{R}^K is known, we prove the following theorem which describes the corresponding secret-message capacity region \mathcal{R}_H^K .

Theorem 3. For $K \leq 3$ or for a symmetric channel with $K > 3$ or for a one-sidedly fair rate tuple $(R_1, \dots, R_K) \in \mathbb{R}_+^K$ with $K > 3$, the secret-message capacity region \mathcal{R}_H^K as defined in Definition 4 is characterized by the following inequality:

$$\max_{j \in \{1, \dots, K\}} \frac{R_j(1 - \frac{\prod_{k=1}^K \delta_k}{\delta_j})}{(1 - \delta_j) \frac{\prod_{k=1}^K \delta_k}{\delta_j} (1 - \prod_{k=1}^K \delta_k)} + \max_{\pi} \sum_{i=1}^K \frac{R_{\pi_i}}{1 - \prod_{k=1}^i \delta_{\pi_k}} \leq 1, \quad (28)$$

where the second maximization is taken over all permutations π of $\{1, \dots, K\}$.

We prove the achievability part of Theorem 3 constructively by describing a linear scheme that achieves any rate tuple in \mathcal{R}_H^K in the mentioned cases. The scheme together with the proof of its properties are given in Section V.

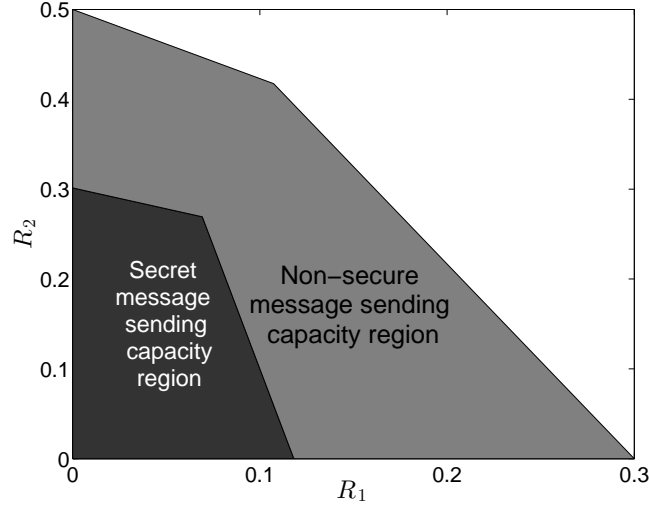


Figure 2. Non-secure message sending and secret-message sending capacity regions for $K = 2$, $\delta_1 = 0.7$, $\delta_2 = 0.5$.

We also develop a converse proof to show that the scheme is optimal. Our converse proof inherently provides a new proof of Theorem 2. We provide the converse proof in Section VII, which completes the proof of Theorem 3.

Comparing regions \mathcal{R}^K and \mathcal{R}_H^K , the first term in (28) can be interpreted as the overhead for security. Indeed, in the scheme that we present, there is a key generation phase whose duration is proportional to this term. In Figure 2, we visualize this overhead for some specific parameter values.

B. Dishonest users

For the case of a dishonest user, we characterize the rate regions $\mathcal{R}_{uDH}^2, \mathcal{R}_{DH}^2$. We focus on security against a dishonest user with side information as defined in Definitions 12-13. In particular, we show that the same rates are achievable against a dishonest user with side information as against honest-but-curious users, *i.e.*, $\mathcal{R}_H^2 = \mathcal{R}_{DH}^2$. This implies $\mathcal{R}_{uDH}^2 = \mathcal{R}_{DH}^2$, hence our result on \mathcal{R}_{DH}^2 implicitly characterizes \mathcal{R}_{uDH}^2 as well. We provide a formal description and proof for $K = 2$, but the same ideas extend for $K > 2$. The following theorem states that $\mathcal{R}_{DH}^2 = \mathcal{R}_H^2$.

Theorem 4. *The rate region \mathcal{R}_{DH}^2 as defined in Definition 13 is the set of all rate pairs $(R_1, R_2) \in \mathbb{R}_+^2$ which satisfy the following two inequalities:*

$$\frac{R_1(1-\delta_2)}{\delta_2(1-\delta_1)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_1\delta_2} \leq 1, \quad (29)$$

$$\frac{R_2(1-\delta_1)}{\delta_1(1-\delta_2)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1\delta_2} + \frac{R_2}{1-\delta_2} \leq 1. \quad (30)$$

It is clear that $\mathcal{R}_{DH}^2 \subseteq \mathcal{R}_H^2$, since the converse developed for the honest-but-curious case provides a valid outer bound. To prove that the region given by (29)-(30) is achievable, we construct a linear scheme that is secure against dishonest users and achieves any pair in the region. The scheme is described in Section VI.

Theorem 4 gives a complete characterization of the problem considering security against a dishonest user with side information. Regarding distribution independent security we do not have such a characterization. We construct a scheme that satisfies this stronger security definition, however its optimality is not clear. The next theorem gives the rate region achieved by our scheme.

Theorem 5. *If a rate pair (R_1, R_2) satisfies*

$$\frac{R_1(1-\delta_2)}{\delta_2(1-\delta_1)(1-\delta_1\delta_2)} + \frac{R_2(1-\delta_1)}{\delta_1(1-\delta_2)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1} + \frac{R_2}{1-\delta_1\delta_2} \leq 1, \quad (31)$$

$$\frac{R_1(1-\delta_2)}{\delta_2(1-\delta_1)(1-\delta_1\delta_2)} + \frac{R_2(1-\delta_1)}{\delta_1(1-\delta_2)(1-\delta_1\delta_2)} + \frac{R_1}{1-\delta_1\delta_2} + \frac{R_2}{1-\delta_2} \leq 1. \quad (32)$$

then $(R_1, R_2) \in \mathcal{R}_{DIS}^2$.

From the definitions it is clear that $\mathcal{R}_{DIS}^2 \subseteq \mathcal{R}_{DH}^2$. We conjecture that there is a fundamental gap between \mathcal{R}_{DIS}^2 and \mathcal{R}_{DH}^2 , and $\mathcal{R}_{DIS}^2 \subset \mathcal{R}_{DH}^2$ holds, but we leave the proof an open question. The scheme that constructively proves Theorem 5 is given in Section VI-D.

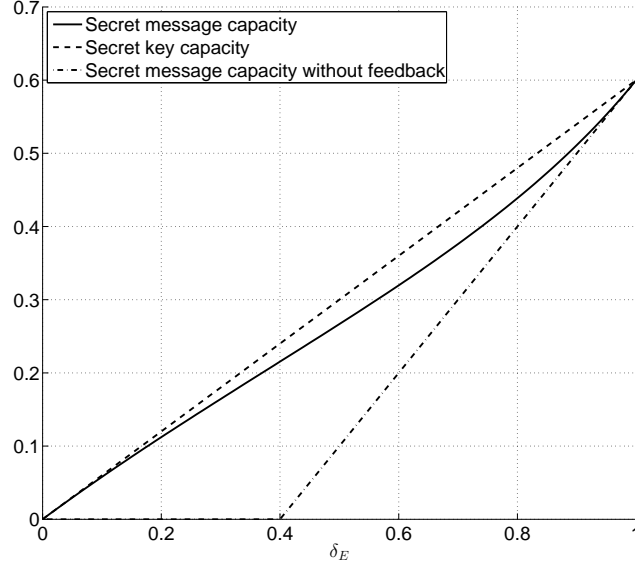


Figure 3. Secret-message and secret-key capacities with and without state-feedback for $\delta = 0.4$.

Security against an eavesdropper: Consider the special case when $K = 2$ and $R_2 = 0$. There is only one receiver with nonzero rate and we aim to secure his message against the other, dishonest party. In this setting the other receiver is equivalent to a passive eavesdropper who overhears the communication. Note that the sender does not trust the feedback from the second receiver, so this feedback is simply ignored. In other words, in this particular setting there is no difference between giving potentially dishonest feedback and not giving any feedback at all. In the end, we have a broadcast channel with one receiver and an eavesdropper against whom we aim to secure a message.

In the light of the argument above, the following definition naturally defines secret-message capacity against an eavesdropper.

Definition 14. *The secret-message capacity \mathcal{C}_E of a broadcast erasure channel with state-feedback against an eavesdropper is the largest R for which $(R, 0) \in \mathcal{R}_{DH}^2$.*

The following corollary characterizes \mathcal{C}_E . The result directly follows from Theorem 4.

Corollary 1. *The secret-message capacity of a broadcast erasure channel with state-feedback against an eavesdropper is*

$$\mathcal{C}_E = (1 - \delta)\delta_E \frac{1 - \delta\delta_E}{1 - \delta\delta_E^2}, \quad (33)$$

where δ_E denotes the erasure probability of the eavesdropper and δ that of the legitimate receiver.

From Wyner's result [14] it is well known that the secret-message capacity of the same setting but without any feedback from the receiver is $(\delta_E - \delta)^+$. Our result shows how feedback helps to increase the achievable rate of secure communication.

Corollary 1 also reveals the subtle difference between the secret-message sending and the secret-key generation problem. It was shown in [7], [24] that in the same setting, a key that is secret from the eavesdropper can be established between the sender and the legitimate receiver at rate $\delta_E(1 - \delta)$, but the key is not known in advance by any of the parties. If we ask that the sender specifies in advance the secret which becomes a shared secret between the sender and the receiver after the protocol run, we arrive to the secret-message sending problem. From (33) it is clear that the rate $\delta_E(1 - \delta)$ is not achievable in this case.

As a comparison, on Figure 3 we plot the secret-message capacity of the broadcast erasure channel against an eavesdropper with and without feedback from the receiver as well as the secret-key capacity of the same setting. Note that without feedback there is no difference between the secret-message sending and the secret-key generation problem.

V. HONEST-BUT-CURIOS USERS

We prove the direct part of Theorem 3 by constructing a secure scheme against honest-but-curious users. At a high level, our scheme consists of two phases:

- 1) *Key generation.* We create K pairwise keys, each key is shared between Alice and one of the receivers, and it is perfectly secure from all the other receivers even if they collude.
- 2) *Encrypted broadcast.* Using the keys set up in the first phase, we employ an encrypted version of the non-secure 1-to- K broadcast scheme as we describe shortly.

Table II
AN EXAMPLE OF THE PROTOCOL RUN.

	Alice sends	Bob's ACK	Calvin's ACK	Bob's key	Calvin's key	Bob decoded	Calvin decoded
Key generation	$\left\{ \begin{array}{l} X_1 \text{ random} \\ X_2 \text{ random} \\ X_3 \text{ random} \end{array} \right.$	$\begin{array}{c} \checkmark \\ \checkmark \\ \times \end{array}$	$\begin{array}{c} \times \\ \checkmark \\ \checkmark \end{array}$	$\begin{array}{c} K_{B,1} = X_1 \\ K_{B,1} \\ K_{B,1} \end{array}$	$K_{C,1} = X_3$		
Encrypted message transmission	$\left\{ \begin{array}{l} X_4 = W_{1,1} \oplus K_{B,1} \\ X_5 = W_{2,1} \oplus K_{C,1} \\ X_6 = W_{2,2} \oplus K_{C,1} \\ X_7 = X_6 \\ X_8 = X_4 \oplus X_7 \end{array} \right.$	$\begin{array}{c} \times \\ \times \\ \times \\ \checkmark \\ \checkmark \end{array}$	$\begin{array}{c} \checkmark \\ \checkmark \\ \times \\ \times \\ \checkmark \end{array}$		$\begin{array}{c} K_{C,1} \\ K_{C,1} \\ K_{C,1} \end{array}$	$\begin{array}{c} \\ \\ \\ W_{1,1} \end{array}$	$\begin{array}{c} W_{2,1} \\ W_{2,1} \\ W_{2,1} \\ W_{2,1}, W_{2,2} \end{array}$

In our second phase, we build on a modified version of the linear scheme presented in [5] that achieves \mathcal{R}^K as stated in Theorem 1. We refer to this scheme as the *non-secure* 1-to- K achievability scheme. Conceptually, this algorithm has two main steps:

Step (a): Alice repeats each message packet $W_{1,1}, \dots, W_{1,N_1}, \dots, W_{K,N_K}$ until at least one of the receivers correctly receives it. We call j the *intended* receiver of a message packet $W_{j,i}$.

Step (b): Alice sends linear combinations of the packets that are not received by their intended receiver in Step (a).

A key contribution of [5] is in specifying how to construct the linear combinations in Step (b) – we refer the reader to [5] for the exact constructions, and highlight here the two important properties that we rely on:

- A message packet successfully delivered to its intended receiver in Step (a) is never used in Step (b).
- The scheme achieves the rate points within \mathcal{R}^K as stated in Theorem 1.

A. Example

Before giving the detailed description of our scheme we show a small example which is suitable to highlight the ideas we use to build our protocol.

Consider a setting with $K = 2$. For convenience, we call the sender Alice, and the two receivers Bob and Calvin. In our example, Alice wants to securely send $N_1 = 1$ message packet $W_1 = [W_{1,1}]$ to Bob and $N_2 = 2$ message packets $W_2 = [W_{2,1}, W_{2,2}]$ to Calvin. The example protocol run is found in Table II.

Key generation:

- (a) Alice transmits random (independent and uniformly distributed) packets X_1, X_2, X_3 . At the end of this phase, Alice and Bob share a secret key packet $K_{B,1} = X_1$ that Bob received and Calvin did not. Similarly, Alice and Calvin share the secret key packet $K_{C,1} = X_3$. The packet X_2 which was received by both Bob and Calvin is discarded.

Encrypted message transmissions:

- (b) Alice secures Bob's first message packet with a one-time pad (using the secret key generated above) and repeatedly transmits an encrypted packet until either Bob or Calvin receives. In our example, this happens immediately (X_4). The packet received only by Calvin is a side information which enables us to efficiently use the channel at a later point.
- (c) In the next few transmissions (X_5 - X_7) we do the same with Calvin's packets. As we see, if only Calvin receives (X_5), a part of the message is successfully delivered, however the key used for encryption can be used again securely to encrypt the next message packet (X_6). If neither Bob nor Calvin receives (X_6), the packet is simply repeated (X_7).
- (d) Once Bob also has a side information (X_7) packet, we send the sum of the two side information packets thereby sending information that is useful simultaneously for both receivers. This happens at transmission $X_8 = X_4 \oplus X_7$, where both Bob and Calvin can decode a novel message packet (X_4 is for Bob, X_7 is for Calvin). Note that at this step we do not need any new keys to secure the transmission.

Through this small example we see the following important features of the scheme:

- The number of key packets we set up and consume is smaller than the number of message packets we convey per user, because we can reuse certain keys if no other receiver has seen any packet encrypted with the given key.
- We exploit side information packets that users have about each other's message to make a single transmission useful for both, without consuming any new key.

B. Detailed description

We need to define a few parameters. The length of the secret keys we aim to set up for receiver j (expressed in terms of packets) is k_j , and the length of the key generation phase in terms of transmissions is n_1 . We define

$$k_j = N_j \frac{1 - \prod_{k=1}^K \delta_k}{1 - \prod_{k=1}^K \delta_k} + \left(N_j \frac{1 - \prod_{k=1}^K \delta_k}{1 - \prod_{k=1}^K \delta_k} \right)^{3/4}, \text{ and } n_1 = \max_j \frac{k_j + k_j^{3/4}}{(1 - \delta_j) \frac{\prod_{k=1}^K \delta_k}{\delta_j}}. \quad (34)$$

1) Key generation

Let K_j denote the key between Alice and receiver j .

Alice transmits n_1 random packets X_1, \dots, X_{n_1} generated uniformly at random over \mathbb{F}_q^L . K_j is the vector of the first k_j packets X_i for which $S_i = j$. If there are less than k_j such packets, we stop and declare an error for receiver j . In other words, Alice transmits random packets, and we treat a packet received by only one receiver as a shared secret between Alice and that receiver.

2) Encrypted broadcast

We now follow the two transmission steps in the non-secure protocol, with the following modifications: in Step (a), we encrypt the message packets using key packets as we specify in the following; in Step (b), we simply reuse the already encrypted packets from Step (a) to create the required linear combinations – we do not use additional key packets.

Step (2.a): Before transmitting each message packet to receiver i , Alice encrypts it by XOR-ing it with a key packet that has either not been used for encryption in the past, or if used, none of the other users received the corresponding packet.

Consider the transmissions to receiver j . Initially, Alice encrypts the first packet for j as $W_{j,1} \oplus K_{j,1}$ and transmits it until it is received by at least one of the receivers. If only receiver j receives this encrypted packet, she reuses the same key packet $K_{j,1}$ to encrypt the next message packet. Subsequently, if for some i and $\ell < N_1$, $k < k_1$: $X_i = W'_{j,\ell} = W_{j,\ell} \oplus K_{j,k}$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } S_i = \emptyset \\ W'_{1,\ell+1} = W_{j,\ell+1} \oplus K_{j,k}, & \text{if } S_i = j \\ W'_{1,\ell+1} = W_{j,\ell+1} \oplus K_{j,k+1}, & \text{otherwise.} \end{cases} \quad (35)$$

In other words, a key is reused until a packet encrypted using it is received by any other receiver. We declare an error if the k_j key packets are not sufficient to encrypt all the N_j message packets of W_j . Alice proceeds similarly for the other keys and messages.

Step (2.b): At the end of Step (2.a), the receivers have received encrypted packets that are not intended for them as side information. We use the same encoding as in Step (b) of the non-secure protocol to deliver these packets to their intended receivers.

C. Analysis of the secure protocol

We need to show that conditions (3)-(6) are all satisfied. Condition (3) is obviously satisfied by construction. We show the other required properties for receiver j , the same arguments apply to any j .

Security: We first argue that our scheme satisfies (6). From construction, at the end of the first phase we create a key K_j with

$$I(K_j; Y_1^{n_1}, \dots, Y_{j-1}^{n_1}, Y_{j+1}^{n_1}, \dots, Y_K^{n_1} S^{n_1}) = 0. \quad (36)$$

In Step (2.a), every packet $W'_{j,\ell}$ that any of the other receivers *receive* has been encrypted using a different key packet $K_{j,i}$; these key packets, from (36), are secret from Calvin and David. Thus the packets received by the $K - 1$ other receivers together are one-time pad encrypted and hence perfectly secret to them, even if they collude. In Step (2.b), Alice transmits linear combinations of packets $W'_{j,\ell}$ that have not been received by receiver j , but have already been received by at least one of the other $K - 1$ receivers. Thus, assuming these receivers collude, they do not receive any innovative $W'_{j,\ell}$. This concludes our argument and shows

$$I(W_j; Y_1^n, \dots, Y_{j-1}^n, Y_{j+1}^n, \dots, Y_K^n S^n) = 0. \quad (37)$$

Decodability: We next prove (4). Trivially, if no error is declared, receiver j can retrieve W_j from W'_j using his key K_j . We show that the probability of declaring an error can be made arbitrarily small. It is enough to consider the following two error events since the other error events are similar: (i) we do not obtain k_j key packets for receiver j during the first phase, and (ii) k_j key packets are not sufficient in Step (2.a).

(i) Let κ denote the number of packets in the first phase that are received only by receiver j . Then, κ is the sum of n_1 i.i.d. Bernoulli variables drawn from $\text{Ber}(p)$, where $p = (1 - \delta_j) \frac{\prod_{k=1}^K \delta_k}{\delta_j}$. Thus,

$$\mathbb{E} \{ \kappa \} = n_1 p = n_1 (1 - \delta_j) \frac{\prod_{k=1}^K \delta_k}{\delta_j} \geq k_j + k_j^{3/4}.$$

The probability of error event (i) equals

$$\Pr \{ \kappa < k_j \} \leq \Pr \{ \mathbb{E} \{ \kappa \} - \kappa > k_j^{3/4} \} \leq \Pr \{ |\mathbb{E} \{ \kappa \} - \kappa| > k_j^{3/4} \} \leq e^{-c\sqrt{k_j}}, \quad (38)$$

for some constant $c > 0$. The last inequality follows from the Chernoff-Hoeffding bound [28]. Selecting N_1 sufficiently large, this error probability can be made arbitrarily small.

(ii) This error event is similar, it occurs if the number of packets that only Bob receives is significantly less than its expected value, and the same technique can be applied to bound the probability of error. With this, we have shown that the scheme is secure against honest-but-curious users as defined in Definition 3.

Rate of the scheme: Finally, as for (5), a straightforward calculation with the given parameters together with the capacity achieving property of non-secure 1-to- K protocol shows that our proposed scheme achieves any rate tuple within the region given by (28). For completeness, we provide the rate calculation in Appendix B. This concludes the proof of the achievability part of Theorem 3.

VI. DISHONEST USERS

We consider the case when $K = 2$ and one of the receivers potentially acknowledges dishonestly. The security of the scheme that we presented in the previous section crucially relies on honest feedback from *all* receivers. If we want to provide security against dishonest users, then the secrecy of message W_j should rely only on the acknowledgment of receiver j . The scheme we describe in this section provides this property.

A. Principles

The structure of the new scheme follows the two-phase structure described previously. However, when we create a key for user j or when we send an encrypted packet to him, instead of the feedback of the other user, we rely on the expected behavior channel. Interestingly, this does not require a sacrifice in rate as long as messages are independent and the message distribution of the dishonest user is uniform.

For illustration, consider the key generation phase. Assume that Alice transmits three random packets X_1, X_2, X_3 , and assume Bob receives X_1, X_2 , while Calvin receives X_2, X_3 as seen in our example in Table II. If we cannot rely on Bob's and Calvin's honesty, but we do know that Bob and Calvin have received at most 2 packets each, we could allocate $K_1 = X_1 \oplus X_2$ as the key between Alice and Bob, and $K_2 = X_2 \oplus X_3$ as the key shared by Alice and Calvin. Note that the number of such linear combinations that we can securely produce is the same as number of key packets that we could set up assuming honest feedback.

We can exploit the channel behavior also in the second phase such that we still have the property that the number of key packets needed is less than the number of message packets to secure. Assume now that Bob has a key $K_1 = [K_{1,1}, K_{1,2}]$. When we send encrypted packets to Bob, assume we expect Calvin to receive two out of three such transmissions – but we do not know which two. We then create three linear combinations of Bob's keys, say $K'_{1,1} = K_{B,1}$, $K'_{1,2} = K_{B,2}$, $K'_{1,3} = K_{B,1} \oplus K_{B,2}$, and transmit $W_{1,1} \oplus K'_{1,1}$, $W_{1,2} \oplus K'_{1,2}$, and $W_{1,3} \oplus K'_{1,3}$. No matter which two of these Calvin receives the message remains secret. Our protocol builds on these ideas.

B. Detailed description

Here we give the details of both phases. We observe that the dishonest user can deny the reception of side information packets by which he can hinder the use of XOR-ed transmissions. The honest user must not experience any decrease in rate even in that case. We limit the length of each step of the scheme to ensure this property.

The operation of the protocol utilizes a set of parameters which can be directly calculated before the protocol starts, and whose use will be described in the following.

$$k_B = N_1 \frac{1 - \delta_2}{1 - \delta_1 \delta_2} + \left(N_1 \frac{1 - \delta_2}{1 - \delta_1 \delta_2} \right)^{3/4} \quad (39)$$

$$k_C = N_2 \frac{1 - \delta_1}{1 - \delta_1 \delta_2} + \left(N_2 \frac{1 - \delta_1}{1 - \delta_1 \delta_2} \right)^{3/4} \quad (40)$$

$$k_1 = \frac{k_B}{\delta_2} + \frac{1}{\delta_2} \left(\frac{2k_B}{\delta_2} \right)^{3/4} \quad k_2 = \frac{k_C}{\delta_1} + \frac{1}{\delta_1} \left(\frac{2k_C}{\delta_1} \right)^{3/4} \quad (41)$$

$$n_1 = \max \left\{ \frac{k_1}{1 - \delta_1} + \left(\frac{k_1}{1 - \delta_1} \right)^{3/4}, \frac{k_2}{1 - \delta_2} + \left(\frac{k_2}{1 - \delta_2} \right)^{3/4} \right\} \quad (42)$$

$$n_{2,1} = \frac{N_1}{1 - \delta_1 \delta_2} + \left(\frac{N_1}{1 - \delta_1 \delta_2} \right)^{3/4} \quad (43)$$

$$n_{2,2} = \frac{N_2}{1 - \delta_1 \delta_2} + \left(\frac{N_2}{1 - \delta_1 \delta_2} \right)^{3/4} \quad (44)$$

$$n'_{2,3} = \frac{N_1}{1-\delta_1} + \left(\frac{N_1}{1-\delta_1}\right)^{3/4} - n_{2,1} \quad (45)$$

$$n''_{2,3} = \frac{N_2}{1-\delta_2} + \left(\frac{N_2}{1-\delta_2}\right)^{3/4} - n_{2,2} \quad (46)$$

$$n = n_1 + n_{2,1} + n_{2,2} + \max\{n'_{2,3}, n''_{2,3}\}. \quad (47)$$

Using our protocol, Alice attempts to send N_1 message packets $W_1 = (W_{1,1}, \dots, W_{1,N_1})$ to Bob and N_2 message packets $W_2 = (W_{2,1}, \dots, W_{2,N_2})$ to Calvin using at most n packet transmissions. We show in Section VI-C that the probability she fails to do so can be made arbitrarily small. She proceeds in two steps.

Key Generation

- 1) Alice transmits n_1 packets X_1, \dots, X_{n_1} generated uniformly at random.
- 2) If Bob receives less than k_1 packets we declare a protocol error for him. Similarly, an error is declared for Calvin if he receives less than k_2 packets. When an error is declared for both users, the protocol terminates. If not, we continue with the user not in error, as if the user in error did not exist.
- 3) Let X_1^B be a $L \times k_1$ matrix that has as columns the first k_1 packets that Bob acknowledged. Alice and Bob create k_B secret key packets as $K_B = X_1^B G_{K_B}$, where G_{K_B} is a $(k_1 \times k_B)$ matrix and is a parity check matrix of a $(k_1, k_1 - k_B)$ maximum distance separable (MDS) code [29]. Similarly, using the first k_2 packets that Calvin acknowledges, Alice and Calvin create k_C secret key packets using the matrix G_{K_C} . Matrices G_{K_B}, G_{K_C} are publicly known and fixed in advance.

Message encryption and transmission

Encryption

- 4) Alice and Bob produce N_1 linear combinations of their k_B secret key packets as $K'_B = K_B G_{K'_B}$, where $G_{K'_B}$ is a $(k_B \times N_1)$ matrix and is a generator matrix of an (N_1, k_B) MDS code which is also publicly known. Similarly, Alice and Calvin create N_2 linear combinations of their k_C key packets.
- 5) Alice creates N_1 encrypted messages to send to Bob

$$U_{B,i} = W_{1,i} \oplus K'_{B,i}, \quad i = 1 \dots N_1, \quad (48)$$

where \oplus is addition in the \mathbb{F}_q^L vector space. Let U_B denote the set of $U_{B,i}, i = 1, \dots, N_1$. She similarly produces a set U_C of N_2 encrypted messages to send to Calvin

$$U_{C,i} = W_{2,i} \oplus K'_{C,i}, \quad i = 1 \dots N_2. \quad (49)$$

Encrypted transmissions

- 6) Alice sequentially takes the first encrypted packet from $U_{B,i}, i = 1 \dots N_1$, that is not yet acknowledged by either Bob or Calvin and repeatedly transmits it, until it is acknowledged by either receiver. That is, if at time i Alice transmits $X_i = U_{B,j}$ for some $j < N_1$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } S_i^* = \emptyset \\ U_{B,j+1}, & \text{otherwise.} \end{cases} \quad (50)$$

Alice continues these transmissions until all packets from U_B are acknowledged or $n_{2,1}$ transmissions are already made in this step. In the former case, she continues with the next step. In the latter case, if Bob does not acknowledge $\frac{N_1(1-\delta_1)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Calvin's packets using ARQ. Similarly, if Calvin does not acknowledge $\frac{N_1(1-\delta_2)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Bob's packets. In case neither receiver is considered to be dishonest, still U_B is not completely delivered, Alice stops and an error is declared for both receivers.

- 7) Similarly, Alice sends not-yet-acknowledged encrypted packets from $U_{C,i}, i = 1 \dots N_2$, until either Bob or Calvin acknowledges. If at time i Alice transmits $X_i = U_{C,j}$ for some $j < N_2$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } S_i^* = \emptyset \\ U_{C,j+1}, & \text{otherwise.} \end{cases} \quad (51)$$

Alice continues these transmissions until all packets from U_C are acknowledged or $n_{2,2}$ transmissions are already made in this step. In the former case, she continues with the next step. In the latter case, if Bob does not acknowledge $\frac{N_2(1-\delta_1)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Calvin's packets using ARQ. Similarly, if Calvin does not acknowledge $\frac{N_2(1-\delta_2)}{1-\delta_1\delta_2}$ packets, then he is considered to be dishonest and Alice continues with sending only Bob's packets. In case neither receiver is considered to be dishonest, still U_C is not completely delivered, Alice stops and an error is declared for both receivers.

- 8) Let Q_B denote the set of packets that only Calvin acknowledged in Step 6. Similarly, Q_C denotes those packets that only Bob acknowledged in Step 7. Alice sequentially takes packets from Q_B and Q_C . For each transmission, she takes

the first packet from Q_B that Bob has not acknowledged together with the first packet from Q_C that Calvin has not yet acknowledged and she transmits the XOR of the two packets. If at time i Alice transmits $X_i = Q_{B,j} \oplus Q_{C,\ell}$ for some $j < |Q_B|, \ell < |Q_C|$, then

$$X_{i+1} = \begin{cases} X_i, & \text{if } S_i^* = \emptyset, \\ Q_{B,j+1} \oplus Q_{C,\ell}, & \text{if } S_i^* = \{1\}, \\ Q_{B,j} \oplus Q_{C,\ell+1}, & \text{if } S_i^* = \{2\}, \\ Q_{B,j+1} \oplus Q_{C,\ell+1}, & \text{if } S_i^* = \{1, 2\}. \end{cases} \quad (52)$$

Alice continues with the XOR-ed transmissions until either receiver acknowledges all his packets. If Bob has already acknowledged all packets from Q_B , Alice repeats packets that are not yet acknowledged by Calvin from Q_C . Similarly, if Calvin has already acknowledged all packets from Q_C , then Alice continues with repeating the remaining packets for Bob from Q_B .

If at any point, the overall number of transmissions would exceed n as defined in (47) we stop and declare an error for the party (or parties) who has not acknowledged all his encrypted message packets.

C. Analysis

Below, we prove that the above scheme is secure against a dishonest user with side information and runs without error with high probability. The rate assertion of the theorem follows from a simple numerical evaluation with the given parameter values.

1) *Security*: In our argument we focus on the secrecy of W_1 against a dishonest Calvin, but the same reasoning works for W_2 against a dishonest Bob as well. Since we do not intend to give security guarantees to a dishonest user and consider at most one user to be dishonest, we may assume that Bob is honest. Moreover, under our definition of dishonest user with side information, W_1 and W_2 are independent and the latter is uniformly distributed over its alphabet, but the distribution of W_1 is arbitrary and controlled by the dishonest Calvin.

To analyze the secrecy of W_1 , we may, without loss of generality, assume that no error was declared for Bob during the key generation phase. Recall that an error is declared for Bob only if Bob fails to acknowledge at least k_1 packets. If an error was in fact declared for Bob, no information about Bob's message W_1 is ever transmitted by Alice. However, note that we do account for this error event when we analyze the probability of error for Bob in the Section VI-C2.

We need to show that the secrecy condition (23) is satisfied by the scheme, *i.e.*, Bob's message remains secret from Calvin even if Calvin controls the distribution of W_1 , and applies any acknowledging strategy. In the proof we omit taking the maximum, but the argument holds for any P_{W_1} and for any adversarial strategy, so the statement follows.

We use the following three lemmas.

Lemma 1. *When Bob is honest and no error is declared for Bob in the key generation phase,*

$$I(K_B; Y_2^{n_1} S^{n_1}) \leq k_B e^{-c_1 \sqrt{k_1}}, \quad (53)$$

if $k_1 = \frac{k_B}{\delta_2} + \frac{1}{\delta_2} \left(\frac{2k_B}{\delta_2} \right)^{3/4}$ and $k_B \geq \frac{2}{\delta_2}$, where $c_1 > 0$ is some constant. Moreover, K_B is uniformly distributed over its alphabet.

Lemma 1 shows that $I(K_B; Y_2^{n_1} S^{n_1})$ can be made small, *i.e.*, the key generation phase is secure. The key facts we use in proving this lemma are (i) the number of packets seen by Calvin concentrates around its mean and (ii) an MDS parity check matrix can be used to perform privacy amplification in the packet erasure setting. The proof is delegated to Appendix A-A.

Let $1_{B,i}^C$ be the indicator random variable for the event that Calvin observes the packet $U_{B,i}$ either in its pure form or in a form where the $U_{B,i}$ packet is added with some $U_{C,j}$ packet. Let M_B^C be the random variable which denotes the number of distinct packets of U_B that Calvin observes, so $M_B^C = \sum_{i=1}^{N_1} 1_{B,i}^C$. Given this notation, we have the following lemmas:

Lemma 2. $H(Y_2^n | Y_2^{n_1} S^n \Theta_2 U_C) \leq \mathbb{E} \{M_B^C\}.$

We prove Lemma 2 in Appendix A-B.

Lemma 3. $H(Y_2^n | W_1 Y_2^{n_1} S^n \Theta_2 U_C) \geq \mathbb{E} \{ \min(k_B, M_B^C) \} - I(K_B; Y_2^{n_1} S^{n_1}).$

We prove Lemma 3 in Appendix A-C.

Using the results of Lemmas 1-3, we conclude the proof as follows. We have that

$$I(W_1; Y_2^n S^n \Theta_2) \leq I(W_1; Y_2^n S^n \Theta_2 U_C) = I(W_1; Y_2^n | Y_2^{n_1} S^n \Theta_2 U_C), \quad (54)$$

where the last equality used the fact that $\Theta_A, \Theta_2, W_2, S^n$ are independent of W_1 and we may express $Y_2^{n_1}, U_C$ as deterministic functions of $\Theta_A, \Theta_2, W_2, S^n$. We use Lemmas 2-3 in (54), to get

$$I(W_1; Y_2^n S^n \Theta_2) \leq \mathbb{E} \{ \max(0, M_B^C - k_B) \} + I(K_B; Y_2^{n_1} S^{n_1}). \quad (55)$$

Lemma 1 gives a bound for the second term. We bound the first term using concentration inequalities, in particular, we use the Chernoff-Hoeffding bound [28] for the purpose. In order to do this, let $Z_{B,i}$ be the number of repetitions of a packet $U_{B,i}$ that Alice makes until Bob acknowledges it (where we count both the transmission in pure form and in addition with some packet from U_C). Note that the random variables $Z_{B,i}$ are independent of each other and have the same distribution. This follows from the fact that the S_i sequence is i.i.d., and each S_i is independent of $(Y_2^{i-1}, S^{i-1}, \Theta_2)$. In other words, Calvin can exert no control over the channel state. Further, for the same reason, with every repetition the chance that Calvin obtains the transmission is $1 - \delta_2$. This implies that the indicator random variables $1_{B,i}^C$ are i.i.d. with

$$\Pr \{1_{B,i}^C = 1\} = (1 - \delta_2) + \delta_1 \delta_2 (1 - \delta_2) + \dots = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}. \quad (56)$$

Notice that M_B^C is a sum of N_1 such independent random variables, and hence $\mathbb{E} \{M_B^C\} = N_1 \frac{1 - \delta_2}{1 - \delta_1 \delta_2}$. Since $k_B = N_1 \frac{1 - \delta_2}{1 - \delta_1 \delta_2} + \left(N_1 \frac{1 - \delta_2}{1 - \delta_1 \delta_2}\right)^{3/4}$, by applying Chernoff-Hoeffding bound we have

$$\mathbb{E} \{\max(0, M_B^C - k_B)\} \leq N_1 \Pr \{M_B^C > k_B\} \leq N_1 e^{-c_2 \sqrt{N_1}}, \quad (57)$$

for a constant $c_2 > 0$. Substituting this back to (55) and using Lemma 1, we get

$$I(W_1; Y_2^n S^n \Theta_2) \leq N_1 e^{-c_2 \sqrt{N_1}} + k_B e^{-c_2 \sqrt{k_B}}, \quad (58)$$

for constants $c_1, c_2 > 0$. By choosing a large enough value of N_1 , we may meet (23)².

2) *Error probability*: We need to bound the probability that an error is declared for Bob. If there is no error for Bob, he will be able to decode W_1 . An error happens if (i) Bob receives less than k_1 packets in the first phase, or (ii) he does not receive enough encrypted message packets in Steps 6 and 8 before the protocol terminates. These error events have the same nature as the error events of our scheme for honest-but-curious users. An error happens if Bob collects significantly fewer packets than he is expected to receive in a particular step. The probability of these events can be made arbitrarily small by applying the same technique as in Section V-C. We omit details to avoid repetitive arguments.

A straightforward computation using the parameters in (40)-(47) shows that the achieved rate region matches the region claimed in Theorem 4. We give the calculation in Appendix B.

3) *Complexity considerations*: It is clear from the analysis in Section VI-C that the length n of the scheme grows as $\max\{O(\log^2(\frac{1}{\epsilon})), O(\frac{1}{\epsilon^4})\}$, where ϵ is the security and probability of error parameter, and ϵ' is the gap parameter associated with the rate. The algorithmic complexity is quadratic in n ; quadratic from the matrix multiplication to produce the key.

D. Distribution independent scheme

In the following we describe a scheme which satisfies distribution independent security as defined in Definition 10. Before that we provide the intuition behind the construction. The protocol in Section VI cannot satisfy distribution independent security, because of the following. Assume Calvin knows his message a-priori and he acknowledges dishonestly in the key generation phase. Then, K_C is constructed of packets that Calvin does not know, but in the second phase, Alice uses this K_C to encrypt his message. If Calvin acknowledges honestly in the second phase, then he learns K_C from U_C , because he already knows W_2 . Since K_C is a linear combination of key generation packets, this way Calvin might learn (in expectation) a $n_1(1 - \delta_2) + k_C$ dimension subspace from the space spanned by the key generation packets. The expected number of key generation packets that either Bob or Calvin receives is $n_1(1 - \delta_1 \delta_2)$ and for a large n , $n_1(1 - \delta_2) + k_C + k_B > n_1(1 - \delta_1 \delta_2)$, hence K_B is not independent of Calvin's observation, which means that Bob's key is not secure.

We can overcome this issue if we modify the key generation phase and make sure that no packet used in generating Calvin's key is contributing Bob's key, thus U_C is conditionally independent of Bob's key given Calvin's observation of the protocol and W_2 . This results in two separate key generation phases, one for Bob and one for Calvin.

1) *Protocol description*: We need two parameters for determining the number of key generation packets.

$$n_{1,1} = \frac{k_1}{1 - \delta_1} + \left(\frac{k_1}{1 - \delta_1}\right)^{3/4}, \quad n_{1,2} = \frac{k_2}{1 - \delta_2} + \left(\frac{k_2}{1 - \delta_2}\right)^{3/4}. \quad (59)$$

We use all other parameters and notations as introduced in Section VI.

Key Generation

- 1) Alice transmits $n_{1,1}$ uniformly random packets $X_1, \dots, X_{n_{1,1}}$ independent of W_1, W_2 . From the first k_1 packets that Bob acknowledges (matrix X_1^B) Alice computes Bob's key as $K_B = X_1^B G_{K_B}$. If Bob does not acknowledge k_1 packets we declare an error for him.
- 2) Alice transmits $n_{1,2}$ uniformly random packets $X_{n_{1,1}+1}, \dots, X_{n_{1,1}+n_{1,2}}$ independent of W_1, W_2 . From the first k_2 packets that Calvin acknowledges out of these $n_{1,2}$ packets (matrix X_1^C), Alice computes Calvin's key as $K_C = X_1^C G_{K_C}$. If Calvin does not acknowledge k_1 packets we declare an error for him.

²Recall from (40)-(47) that by saying that we choose N_1 large enough we cause n to be large enough.

- 3) When an error is declared for both users, the protocol terminates. If not, we continue with the user not in error, as if the user in error did not exist.

Message encryption and transmission

Steps 4-8 are exactly the same as described in Section VI.

This scheme provides distribution independent security, which property is proved in Appendix C. The proof follows the same lines as the analysis of Section VI-C. This, together with a straightforward rate calculation completes the proof of Theorem 5.

VII. CONVERSE

We show the converse part of Theorem 3, by which we conclude the proof of Theorems 3, 4 and Corollary 1. This result proves the optimality of the schemes presented in Section V and in Section VI.

Proof: We present our proof for $K = 3$, the generalization of the same argument for any K is straightforward. We are going to show that for any j and any π

$$\frac{R_j(1 - \frac{\delta_1\delta_2\delta_3}{\delta_j})}{(1 - \delta_j)\frac{\delta_1\delta_2\delta_3}{\delta_j}(1 - \delta_1\delta_2\delta_3)} + \frac{R_{\pi_1}}{1 - \delta_{\pi_1}} + \frac{R_{\pi_2}}{1 - \delta_{\pi_1}\delta_{\pi_2}} + \frac{R_{\pi_3}}{1 - \delta_{\pi_1}\delta_{\pi_2}\delta_{\pi_3}} \leq 1 \quad (60)$$

holds, which implies the statement of the theorem. Also, to avoid cumbersome notation we show (60) for $j = 1$ and $\pi = (1, 2, 3)$. With simple relabeling, the same argument holds for any j and π .

$$n \geq \sum_{i=1}^n H(X_i) \geq \sum_{i=1}^n H(X_i|Y_1^{i-1}S^{i-1}) = \sum_{i=1}^n H(X_i|Y_1^{i-1}Y_2^{i-1}S^{i-1}) + I(X_i; Y_2^{i-1}|Y_1^{i-1}S^{i-1}) \quad (61)$$

$$= \sum_{i=1}^n H(X_i|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}S^{i-1}) + I(X_i; Y_2^{i-1}|Y_1^{i-1}S^{i-1}) + I(X_i; Y_3^{i-1}|Y_1^{i-1}Y_2^{i-1}S^{i-1}) \quad (62)$$

$$= \sum_{i=1}^n H(X_i|W_1W_2W_3Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}S^{i-1}) \quad (63)$$

$$+ I(X_i; Y_2^{i-1}|Y_1^{i-1}S^{i-1}) \quad (64)$$

$$+ I(X_i; Y_3^{i-1}|Y_1^{i-1}Y_2^{i-1}S^{i-1}) \quad (65)$$

$$+ I(X_i; W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}S^{i-1}) \quad (66)$$

In the following Lemmas 4-7 we give bounds on each of the terms (63)-(66). Combining these results together gives (60) and in turn the statement of the theorem. The detailed proofs of these lemmas are delegated to Appendix D. ■

A. Proof of Theorem 2

Proof: It is sufficient to prove the inequality for $\pi = (1, 2, 3)$. By relabeling, the same argument holds for any π . By repeating the first steps of the previous proof and bounding term (63) by 0, we have

$$n \geq \sum_{i=1}^n I(X_i; Y_2^{i-1}|Y_1^{i-1}S^{i-1}) \quad (67)$$

$$+ I(X_i; Y_3^{i-1}|Y_1^{i-1}Y_2^{i-1}S^{i-1}) \quad (68)$$

$$+ I(X_i; W_1W_2W_3|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}S^{i-1}) \quad (69)$$

Lemmas 5-7 give bounds on terms (67)-(69) respectively. Combining these gives the stated inequality. ■

B. Interpretation of the converse proof

To facilitate understanding, beside our formal proof through Lemmas 4-10 here we provide some intuitive interpretation of terms (63)-(66) and of the inequalities we derive. It will be helpful to match terms to the steps of our scheme, but we stress that the proof holds for any possible scheme.

In Lemma 4 we see the following (here we omit small terms for simplicity):

$$(1 - \delta_1)\delta_2\delta_3 \sum_{i=1}^n H(X_i|Y_1^{i-1}Y_2^{i-1}Y_3^{i-1}W_1W_2W_3S^{i-1}) \geq \frac{nR_1(1 - \delta_2\delta_3)}{1 - \delta_1\delta_2\delta_3}. \quad (70)$$

The entropy term on the LHS of this inequality accounts for fresh randomness sent by the source. In our scheme we call this the key generation phase. The constant factor $(1 - \delta_1)\delta_2\delta_3$ suggests that a random packet becomes a key for receiver 1 if only he receives the transmission. The RHS of the inequality corresponds to the expected number of (encrypted) W_1 packets that not only receiver 1 gets, but some other receivers also overhear. These are the packets that need to be secured, thus for perfect

secrecy, receiver 1 needs at least the same number of key packets. This lower bound on term (63) suggests that any scheme has to introduce some source randomness. We find it natural to call it key generation.

Terms (64)-(66) correspond to the second phase of our protocol. Term (66) corresponds to the first step of the message transmission phase (see Step (a)), when the sender ensures that the receivers *together* could decode all the messages. Terms (64)-(65) account for the encoded transmissions. E.g. (64) intuitively corresponds to “a packet that is of interest for receiver 1 and known by receiver 2”. Indeed, Lemma 5 lower bounds this term with the expected number of transmissions that are needed to convey to receiver 1 the side information overheard by receiver 2.

C. Lemmas

Here we state the lemmas that we use in the converse proof. The proofs of the lemmas are found in Appendix D.

Lemma 4. *From conditions (3)-(7) it follows that*

$$\sum_{i=1}^n H(X_i | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} W_1 W_2 W_3 S^{i-1}) \geq \frac{nR_1(1 - \delta_2\delta_3)}{(1 - \delta_1)\delta_2\delta_3(1 - \delta_1\delta_2\delta_3)} - \mathcal{E}_8, \quad (71)$$

where $\mathcal{E}_8 = \mathcal{E}_7 \frac{1 - \delta_2\delta_3}{(1 - \delta_1)\delta_2\delta_3}$, and \mathcal{E}_7 is an error constant specified in Lemma 10.

Lemma 5. *From conditions (3)-(7) it follows that*

$$\sum_{i=1}^n I(X_i; Y_2^{i-1} | Y_1^{i-1} S^{i-1}) \geq \frac{nR_1}{1 - \delta_1} - \frac{nR_1}{1 - \delta_1\delta_2} - \mathcal{E}_1, \quad (72)$$

where $\mathcal{E}_1 = \frac{h_2(\epsilon) + \epsilon}{1 - \delta_1}$.

Lemma 6. *From conditions (3)-(7) it follows that*

$$\sum_{i=1}^n I(X_i; Y_3^{i-1} | Y_1^{i-1} Y_2^{i-1} S^{i-1}) \geq \frac{n(R_1 + R_2)}{1 - \delta_1\delta_2} - \frac{n(R_1 + R_2)}{1 - \delta_1\delta_2\delta_3} - \mathcal{E}_2, \quad (73)$$

where $\mathcal{E}_2 = \frac{h_2(2\epsilon) + 2\epsilon}{1 - \delta_1\delta_2}$.

Lemma 7. *From conditions (3)-(7) it follows that*

$$\frac{n(R_1 + R_2 + R_3)}{1 - \delta_1\delta_2\delta_3} - \mathcal{E}_3 \leq \sum_{i=1}^n I(X_i; W_1 W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}) \leq \frac{n(R_1 + R_2 + R_3)}{1 - \delta_1\delta_2\delta_3} \quad (74)$$

where $\mathcal{E}_3 = \frac{h_2(3\epsilon) + 3\epsilon}{1 - \delta_1\delta_2\delta_3}$.

Lemma 8. *From the definition of the channel it follows that*

$$\sum_{i=1}^n H(X_i | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} W_1 W_2 W_3 S^{i-1}) \geq \frac{1 - \delta_2\delta_3}{(1 - \delta_1)\delta_2\delta_3} \sum_{i=1}^n I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} W_1 W_2 W_3 S^{i-1}) \quad (75)$$

Lemma 9. *From the security condition (6) it follows that*

$$\mathcal{E}_4 > \sum_{i=1}^n I(X_i; W_1 | Y_2^{i-1} Y_3^{i-1} S^{i-1}), \quad (76)$$

where $\mathcal{E}_4 = \frac{\epsilon}{1 - \delta_2\delta_3}$.

Lemma 10. *From conditions (3)-(6) it follows that*

$$\sum_{i=1}^n I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1 W_2 W_3) \geq \frac{nR_1}{1 - \delta_1\delta_2\delta_3} - \mathcal{E}_7, \quad (77)$$

where $\mathcal{E}_7 = 2\mathcal{E}'_2 + \mathcal{E}_4 + \mathcal{E}_5 + \mathcal{E}_6$, $\mathcal{E}_5 = \frac{h_2(\epsilon) + \epsilon}{1 - \delta_2\delta_3}$, $\mathcal{E}_6 = \frac{h_2(\epsilon) + \epsilon}{1 - \delta_1\delta_2\delta_3}$, and $\mathcal{E}'_2 = \frac{h_2(2\epsilon) + 2\epsilon}{1 - \delta_2\delta_3}$.

VIII. DISCUSSION

A. Channels with correlated erasures

Our results can be generalized for memoryless channels with arbitrary correlation between the erasure events. We consider the case of honest-but-curious users. Let $\delta_{\mathcal{N}}$ denote the erasure probability that the set \mathcal{N} of receivers jointly experience and p_j the probability that only user j receives:

$$\delta_{\mathcal{N}} = \Pr \{ \forall j \in \mathcal{N} : Y_{j,i} = \perp \} \quad (78)$$

$$p_j = \Pr \{ Y_{j,i} = X_{j,i}, \forall k \neq j : Y_{k,i} = \perp \}. \quad (79)$$

Our outer bound proof relies on knowing the statistical behavior of the channel but not on its independence property. Using the parameters defined above, a straightforward generalization of the proof in Section VII results in a more general bound: any rate tuple $(R_1, \dots, R_K) \in \mathbb{R}_+^K$ in \mathcal{R}_H^K satisfies

$$\max_{j \in \{1, \dots, K\}} \frac{R_j(1 - \delta_{\{-j\}})}{p_j(1 - \delta_{\{1, \dots, K\}})} + \max_{\pi} \sum_{i=1}^K \frac{R_{\pi_i}}{1 - \delta_{\{\pi_1, \dots, \pi_i\}}} \leq 1, \quad (80)$$

where $\delta_{\{-j\}}$ is used as a shorthand for $\delta_{\{1, \dots, j-1, j+1, \dots, K\}}$.

Our arguments for the key generation phase also do not exploit the independence property of the channel. It follows that for the broadcast erasure channel the upper bound on the achievable key generation rate for user j

$$\max I(X_i; Y_{j,i} | Y_{1,i}, \dots, Y_{j-1,i}, Y_{j+1,i}, \dots, Y_{K,i}) = p_j \quad (81)$$

derived in [7] is achievable without the requirement of independent erasures. In the parameter definitions and proofs p_j is substituted as the key generation rate for user j . Consider our parameter definitions for honest-but-curious users in eq. (34). With this modification,

$$k_j = N_j \frac{1 - \delta_{\{-j\}}}{1 - \delta_{\{1, \dots, K\}}} + \left(N_j \frac{1 - \delta_{\{-j\}}}{1 - \delta_{\{1, \dots, K\}}} \right)^{3/4}, \text{ and } n_1 = \max_j \frac{k_j + k_j^{3/4}}{p_j}. \quad (82)$$

The second phase of our achievability algorithm (Step (2.b)) uses a capacity achieving non-secure coding scheme. For the cases where such a scheme is available (for non-secure coding schemes we refer the reader to [5]), our secure protocol naturally extends for a channel with correlated erasures.

B. Security notions

We formulate our results in information theoretic terms, defining secrecy as a mutual information term being negligibly small. In the realm of computational cryptography it is more common to prove security of an encryption scheme by showing distinguishing security or semantic security. To facilitate the interpretation of our results and to allow a fair comparison with other schemes, we cite a recent result from [13], which shows equivalence between the two approaches. However, the definitions of distinguishing security and semantic security are applicable only for a single user setting. We can directly use these definitions for the special cases when $R_1 = 0$ or $R_2 = 0$, i.e., when we consider security against an eavesdropper. In these cases, our definition of security against a dishonest user with side information is equivalent to semantic security. We extend the notion of semantic security such that it handles joint message distributions, which results in a definition matching distribution independent security. We will give the definitions for Bob's security, the security for Calvin is completely symmetric.

It is common to define the *advantage* of the adversary to express the gain that the adversary obtains by observing the protocol. Considering security against an eavesdropper, the adversarial advantage expressed in terms of mutual information (mis = mutual information security) is defined as:

$$\mathbf{Adv}^{\text{mis}} = \max_{P_{W_1}, \sigma_2} I(W_1; Y_2^n S^n \Theta_2). \quad (83)$$

The notion of semantic security captures the intuition that the probability that an adversary can compute a function f of the message should not increase significantly after observing the protocol compared to the a-priori probability of a correct guess. The semantic security advantage is defined as

$$\mathbf{Adv}^{\text{ss}} = \max_{f, P_{W_1}, \sigma_2} \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \sigma_2) = f(W_1) \} - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f) = f(W_1) \} \right\}, \quad (84)$$

where f is an arbitrary function of W_1 , \mathcal{A} is any function that the adversary may compute after observing the protocol and \mathcal{S} is a simulator trying to compute f without accessing the protocol output. Here also, W_2 is uniformly distributed and independent of W_1 . The term simulator to denote guessing functions comes from the intuition that ideally there exists an algorithm (simulator)

that simulates the run of a protocol without having access to the message and whose output is indistinguishable from the output of a real protocol. Theorems 1, 5 and 8 from [13] prove the following inequalities:

$$\mathbf{Adv}^{\text{ss}} \leq \sqrt{2 \cdot \mathbf{Adv}^{\text{mis}}}; \quad \mathbf{Adv}^{\text{mis}} \leq 4 \cdot \mathbf{Adv}^{\text{ss}} \log \left(\frac{2^n}{\mathbf{Adv}^{\text{ss}}} \right) \quad (85)$$

This result shows that requirement (23) is naturally equivalent to semantic security, *i.e.*, a small ϵ in (23) implies that \mathbf{Adv}^{ss} is also small.

Applying the above definition for a case when $R_2 > 0$ implicitly assumes that Calvin cannot choose the distribution of his own message W_2 . We now extend the definition of semantic security such that it does not rely on the distribution of W_2 , which results in a stronger notion of security. We define the adversarial advantage for this case as

$$\mathbf{Adv}_{\text{dis}}^{\text{ss}} = \max_{f, P_{W_1, W_2}, \sigma_2} \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \sigma_2, W_2) = f(W_1, W_2) \} - \max_{\mathcal{S}} \Pr \{ \mathcal{S}(P_{W_1, W_2}, f, W_2) = f(W_1, W_2) \} \right\}. \quad (86)$$

Note that here we allow the simulator to have access to the message W_2 which an honest Calvin will learn. The corresponding definition of adversarial advantage for distribution independent security directly comes from (18):

$$\mathbf{Adv}_{\text{dis}}^{\text{mis}} = \max_{P_{W_1, W_2}, \sigma_2} I(W_1; Y_2^n S^n \Theta_2 | W_2) \quad (87)$$

We show in Appendix E the following lemma, which implies that requirement (18) is equivalent to this extended notion of semantic security.

Lemma 11.

$$\mathbf{Adv}_{\text{dis}}^{\text{ss}} \leq \sqrt{2 \cdot \mathbf{Adv}_{\text{dis}}^{\text{mis}}} \quad (88)$$

$$\mathbf{Adv}_{\text{dis}}^{\text{mis}} \leq 4 \cdot \mathbf{Adv}_{\text{dis}}^{\text{ss}} \log \left(\frac{2^n}{\mathbf{Adv}_{\text{dis}}^{\text{ss}}} \right). \quad (89)$$

These results show that although security definitions might look quite different at first sight, there is no fundamental difference between these notions of security. As a corollary, our results also characterize the rate regions that are achievable by any scheme that provides semantic security.

REFERENCES

- [1] L. Czap, V. Prabhakaran, C. Fragouli, and S. Diggavi, "Secret Message Capacity of Erasure Broadcast Channels with Feedback," in *Information Theory Workshop (ITW)*, 2011, pp. 65–69.
- [2] L. Czap, V. Prabhakaran, S. Diggavi, and C. Fragouli, "Broadcasting Private Messages Securely," in *International Symposium on Information Theory (ISIT)*. IEEE, 2012, pp. 428–432.
- [3] L. Czap, V. M. Prabhakaran, S. Diggavi, and C. Fragouli, "Securing Broadcast Against Dishonest Receivers," in *International Symposium on Network Coding (NetCod)*, 2013.
- [4] M. Siavoshani, S. Mishra, S. Diggavi, and C. Fragouli, "Group Secret Key Agreement over State-dependent Wireless Broadcast Channels," in *International Symposium on Information Theory (ISIT)*, 2011, pp. 1960–1964.
- [5] C. Wang, "On the Capacity of 1-to-K Broadcast Packet Erasure Channels with Channel Output Feedback," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 931–956, 2012.
- [6] M. Gatzianas, L. Georgiadis, and L. Tassioulas, "Multiuser Broadcast Erasure Channel with Feedback – Capacity and Algorithms," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5779–5804, Sept 2013.
- [7] U. Maurer, "Secret Key Agreement by Public Discussion from Common Information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
- [8] M. J. Siavoshani, U. Pulletti, E. Atsan, I. Safaka, C. Fragouli, K. Argyraki, and S. Diggavi, "Exchanging Secrets Without Using Cryptography," ArXiv, abs/1105.4991v1, May 2011. [Online]. Available: <http://arxiv.org/abs/1105.4991v1>; <http://arxiv.org/pdf/1105.4991v1>
- [9] Y. Abdallah, M. A. Latif, M. Youssef, A. Sultan, and H. E. Gamal, "Keys Through ARQ: Theory and Practice," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 737–751, September 2011.
- [10] N. Chandran, B. Kanukurthi, R. Ostrovsky, and L. Reyzin, "Privacy Amplification with Asymptotically Optimal Entropy Loss," in *42nd ACM symposium on Theory of computing*. ACM, 2010, pp. 785–794.
- [11] I. Safaka, C. Fragouli, K. Argyraki, and S. Diggavi, "Exchanging pairwise secrets efficiently," in *INFOCOM*, April 2013, pp. 2265–2273.
- [12] K. Argyraki, S. Diggavi, M. Duarte, C. Fragouli, M. Gatzianas, and P. Kostopoulos, "Creating Secrets out of Erasures," in *Conference on Mobile Computing and Networking (MobiCom)*. New York, NY, USA: ACM, 2013, pp. 429–440.
- [13] M. Bellare, S. Tessaro, and A. Vardy, "Semantic Security for the Wiretap Channel," in *International Cryptology Conference (CRYPTO)*. Springer, 2012, pp. 294–311.
- [14] A. D. Wyner, "The Wire-tap Channel," *The Bell system Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.
- [15] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.
- [16] Y. Liang, H. V. Poor, and S. Shamai, "Information Theoretic Security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2009.
- [17] R. Ahlswede and I. Csiszár, "Common Randomness in Information Theory and Cryptography - I: Secret Sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.
- [18] I. Csiszár and P. Narayan, "Secrecy Capacities for Multiple Terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, 2004.
- [19] —, "Secrecy Capacities for Multiterminal Channels," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 2437–2452, 2008.
- [20] R. Ahlswede and N. Cai, *Transmission, Identification and Common Randomness Capacities for Wire-Tape Channels with Secure Feedback from the Decoder.*, ser. LNCS. Springer, 2006, vol. 4123.

- [21] L. Lai, H. E. Gamal, and H. Poor, "The Wiretap Channel with Feedback: Encryption over the Channel," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5059–5067, 2008.
- [22] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y. Kim, "Wiretap Channel with Secure Rate-limited Feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5353–5361, 2009.
- [23] M. Jafari Siavoshani, S. Mishra, S. Diggavi, and C. Fragouli, "Group Secret Key Agreement over State-dependent Wireless Broadcast Channels," in *IEEE International Symposium on Information Theory (ISIT)*, 2011.
- [24] M. Jafari Siavoshani, S. Diggavi, C. Fragouli, U. K. Pulleti, and K. Argyraki, "Group Secret Key Generation over Broadcast Erasure Channels," in *Asilomar Conference on Signals, Systems, and Computers*, 2010, pp. 719–723.
- [25] R. Yeung, "Secure Network Coding," in *Proceedings IEEE International Symposium on Information Theory*, Ieee, 2005, p. 323.
- [26] L. Georgiadis and L. Tassiulas, "Broadcast Erasure Channel with Feedback-capacity and Algorithms," in *Workshop on Network Coding, Theory, and Applications, (NetCod)*. IEEE, 2009, pp. 54–61.
- [27] M. Maddah-Ali and D. Tse, "Completely Stale Transmitter Channel State Information is Still Very Useful," in *Annual Allerton Conference on Communication, Control, and Computing*, 2010, pp. 1188–1195.
- [28] W. Hoeffding, "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American statistical association*, vol. 58, no. 301, pp. 13–30, 1963.
- [29] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes*, 2nd ed. North-holland Publishing Company, 1978.

APPENDIX A

PROOF OF LEMMAS IN SECTION VI-C

A. Proof of Lemma 1

Proof: With a slight abuse of notation, in the following X_1^{BC} will denote the *actual* packets Calvin received (not necessarily the same as those that he acknowledges) out of the first k_1 packets that Bob received. Note that here we assume that an error was not declared for Bob in the key generation phase and hence Bob did receive at least k_1 packets in the key generation phase. Also let $X_1^{B\emptyset}$ be the packets seen only by Bob among the first k_1 he receives. Let $I_{B\emptyset}$ and I_{BC} be the index sets corresponding to $X_1^{B\emptyset}$ and X_1^{BC} . Recall that X_1^B denotes the first k_1 packets received by Bob. The notation M^I will denote a matrix M restricted to the columns defined by index set I . Given this,

$$I(K_B; Y_2^{n_1} S^{n_1}) = I(X_1^B G_{K_B}; X_1^{BC} S^n) \quad (90)$$

$$= H(X_1^B G_{K_B}) - H(X_1^B G_{K_B} | X_1^{BC} S^n) \quad (91)$$

$$= k_B - H(X_1^B G_{K_B} | X_1^{BC} S^n) \quad (92)$$

$$= k_B - H\left[\begin{bmatrix} X_1^{B\emptyset} G_{K_B}^{I_{B\emptyset}} & X_1^{BC} G_{K_B}^{I_{BC}} \end{bmatrix} | X_1^{BC} S^n\right] \quad (93)$$

$$= k_B - H(X_1^{B\emptyset} G_{K_B}^{I_{B\emptyset}} | X_1^{BC} S^n) \quad (94)$$

$$= k_B - H(X_1^{B\emptyset} G_{K_B}^{I_{B\emptyset}} | S^n), \quad (95)$$

where the third equality follows from the MDS property of the matrix G_{K_B} . Using the same property, we have

$$H(X_1^{B\emptyset} G_{K_B}^{I_{B\emptyset}} | S^n) = \sum_{i=0}^{k_1} \min\{i, k_B\} L \log q \Pr\{|X_1^{B\emptyset}| = i\} \quad (96)$$

$$\geq k_B \sum_{i=k_B}^{k_1} \Pr\{|X_1^{B\emptyset}| = i\} \quad (97)$$

$$= k_B \Pr\{|X_1^{B\emptyset}| \geq k_B\} \quad (98)$$

$$= k_B \left(1 - \Pr\{|X_1^{B\emptyset}| < k_B\}\right) \quad (99)$$

$$= k_B \left(1 - \Pr\{|X_1^{BC}| \geq k_1 - k_B\}\right) \quad (100)$$

$$\stackrel{(a)}{\geq} k_B \left(1 - \Pr\{|X_1^{BC}| \geq (1 - \delta_2)k_1 + k_1^{3/4}\}\right) \quad (101)$$

$$\geq k_B \left(1 - \Pr\{||X_1^{BC}| - E[|X_1^{BC}|]| > k_1^{3/4}\}\right), \quad (102)$$

where the inequality (a) follows from the fact that the conditions on k_B and k_1 imply that

$$k_1 - k_B \geq (1 - \delta_2)k_1 + k_1^{3/4}.$$

The Chernoff-Hoeffding bound gives that for some constant $c_1 > 0$

$$\Pr\{||X_1^{BC}| - E[|X_1^{BC}|]| > k_1^{3/4}\} \leq e^{-c_1 \sqrt{k_1}}. \quad (103)$$

So, we have that

$$I(K_B; Y^{n_1} S^n) \leq k_B e^{-c_1 \sqrt{k_1}}. \quad (104)$$

The final assertion of the lemma is a simple consequence of the MDS property of the code and the fact that X^{n_1} are uniformly i.i.d. ■

B. Proof of Lemma 2

Proof: Let U_B^C be a vector of length N_1 such that the i -th element $U_{B,i}^C$ is $U_{B,i}$ if Calvin observes this $U_{B,i}$ either in the pure form or added with some element of U_C , and $U_{B,i}^C = \perp$ otherwise. Let $1_{B,i}^C$ is the indicator random variable for the event $U_{B,i}^C \neq \perp$, so $M_B^C = \sum_{i=1}^{N_1} 1_{B,i}^C$. It is easy to see that the following are information equivalent (i.e., we can express each side as a deterministic function of the other)

$$(Y_2^n, S^n, \Theta_2, U_C) \equiv (U_B^C, Y_2^{n_1}, S^n, \Theta_2, U_C).$$

Therefore,

$$H(Y_2^n S^n \Theta_2 U_C) = H(U_B^C Y_2^{n_1} S^n \Theta_2 U_C).$$

$$H(Y_2^n | Y_2^{n_1} S^n \Theta_2 U_C) = H(U_B^C | Y_2^{n_1} S^n \Theta_2 U_C) \quad (105)$$

$$= \sum_{i=1}^{N_1} H(U_{B,i}^C | U_B^{C^{i-1}} Y_2^{n_1} S^n \Theta_2 U_C) \quad (106)$$

$$= \sum_{i=1}^{N_1} H(U_{B,i}^C | 1_{B,i}^C U_B^{C^{i-1}} Y_2^{n_1} S^n \Theta_2 U_C) \quad (107)$$

$$\leq \sum_{i=1}^{N_1} H(U_{B,i}^C | 1_{B,i}^C) \quad (108)$$

$$\leq \sum_{i=1}^{N_1} \Pr \{1_{B,i}^C = 1\} = \mathbb{E} \left\{ \sum_{i=1}^{N_1} 1_{B,i}^C \right\}. \quad (109)$$

where the third equality follows from the fact that the indicator random variable $1_{B,i}^C$ is a deterministic function of the conditioning random variables. ■

C. Proof of Lemma 3

Proof: We adopt the notation for U_B^C and $1_{B,i}^C$ introduced in the proof of Lemma 2. In addition, let K_B^C be defined in a similar manner as U_B^C such that $K_{B,i}^C = \perp$ if $U_{B,i}^C = \perp$ and $K_{B,i}^C = K'_{B,i}$ otherwise. Also, let 1_B^C be the vector of indicator random variables $1_{B,i}^C$, $j = 1, \dots, N_1$.

Proceeding as in the proof of Lemma 2, we have

$$H(Y_2^n | W_1 Y_2^{n_1} S^n \Theta_2 U_C) = H(U_B^C | W_1 Y_2^{n_1} S^n \Theta_2 U_C) \quad (110)$$

$$= H(K_B^C | W_1 Y_2^{n_1} S^n \Theta_2 U_C) \geq H(K_B^C | 1_B^C W_1 Y_2^{n_1} S^n \Theta_2 U_C) \quad (111)$$

$$= H(K_B^C | 1_B^C) - I(K_B^C; W_1 Y_2^{n_1} S^n \Theta_2 U_C | 1_B^C) \quad (112)$$

But, from the MDS property of $G_{K_B^C}$, and the fact that K_B is uniformly distributed over its alphabet, we have

$$H(K_B^C | 1_B^C) = \sum_{i=1}^{N_1} \min(i, k_B) \Pr \left\{ \sum_{j=1}^{N_1} 1_{B,j}^C = i \right\} = \mathbb{E} \left\{ \min \left(k_B, \sum_{i=1}^{N_1} 1_{B,i}^C \right) \right\}. \quad (113)$$

Also,

$$I(K_B^C; W_1 Y_2^{n_1} S^n \Theta_2 U_C | 1_B^C) \stackrel{(a)}{=} I(K_B^C; Y_2^{n_1} S^{n_1} | 1_B^C) \leq I(K_B^C 1_B^C; Y_2^{n_1} S^{n_1}) \leq I(K_B; Y_2^{n_1} S^{n_1}). \quad (114)$$

where (a) follows from the fact that the distribution of W_2 (uniform and independent of S^n, Θ_A, Θ_2) implies that U_C is independent of Θ_A, S^n and using this we can argue that the following is Markov chain

$$K_B^C - (1_B^C, Y_2^{n_1}, S^{n_1}) - (W_1, \Theta_2, U_C).$$

By substituting back we obtain the claim of the lemma. ■

APPENDIX B RATE CALCULATION

A. Honest-but-curious users

The rate achieved for user j is $R_j = \lim_{n \rightarrow \infty} \frac{N_j}{n}$. Compared to the non-secure 1-to- K protocol we have an overhead of n_1 transmissions. We have

$$\lim_{n \rightarrow \infty} \frac{k_j}{n} = R_j \frac{1 - \frac{\prod_{k=1}^K \delta_k}{\delta_j}}{1 - \prod_{k=1}^K \delta_k}, \quad (115)$$

and thus

$$\lim_{n \rightarrow \infty} \frac{k_j + k_j^{3/4}}{n} = R_j \frac{1 - \frac{\prod_{k=1}^K \delta_k}{\delta_j}}{1 - \prod_{k=1}^K \delta_k}, \quad (116)$$

$$\lim_{n \rightarrow \infty} \frac{n_1}{n} = \max_{j \in \{1, \dots, K\}} \frac{R_j (1 - \frac{\prod_{k=1}^K \delta_k}{\delta_j})}{(1 - \delta_j) \frac{\prod_{k=1}^K \delta_k}{\delta_j} (1 - \prod_{k=1}^K \delta_k)}. \quad (117)$$

Using Theorem 1 the rate assertion of Theorem 3 follows.

B. Dishonest user

Similarly as in the honest-but-curious case, we need to compute $\lim_{n \rightarrow \infty} \frac{n_1}{n}$ and $\lim_{n \rightarrow \infty} \frac{\max\{n'_2, n''_2\}}{n}$. It is immediate that

$$\lim_{n \rightarrow \infty} \frac{n'_2}{n} = \frac{R_1}{1 - \delta_1} + \frac{R_2}{1 - \delta_1 \delta_2} \quad (118)$$

$$\lim_{n \rightarrow \infty} \frac{n''_2}{n} = \frac{R_1}{1 - \delta_1 \delta_2} + \frac{R_2}{1 - \delta_2} \quad (119)$$

Further

$$\lim_{n \rightarrow \infty} \frac{k_B}{n} = R_1 \frac{1 - \delta_2}{1 - \delta_1 \delta_2} \quad (120)$$

$$\lim_{n \rightarrow \infty} \frac{k_C}{n} = R_2 \frac{1 - \delta_1}{1 - \delta_1 \delta_2}, \quad (121)$$

from which

$$\lim_{n \rightarrow \infty} \frac{k_1}{n} = R_1 \frac{1 - \delta_2}{\delta_2 (1 - \delta_1 \delta_2)} \quad (122)$$

$$\lim_{n \rightarrow \infty} \frac{k_2}{n} = R_2 \frac{1 - \delta_1}{\delta_1 (1 - \delta_1 \delta_2)}, \quad (123)$$

and

$$\lim_{n \rightarrow \infty} \frac{n_1}{n} = \max \left(R_1 \frac{1 - \delta_2}{\delta_2 (1 - \delta_1) (1 - \delta_1 \delta_2)}, R_2 \frac{1 - \delta_1}{\delta_1 (1 - \delta_2) (1 - \delta_1 \delta_2)} \right). \quad (124)$$

We also observe that

$$R_1 \frac{1 - \delta_2}{\delta_2 (1 - \delta_1) (1 - \delta_1 \delta_2)} > R_2 \frac{1 - \delta_1}{\delta_1 (1 - \delta_2) (1 - \delta_1 \delta_2)} \Leftrightarrow \frac{R_1}{1 - \delta_1} + \frac{R_2}{1 - \delta_1 \delta_2} > \frac{R_1}{1 - \delta_1 \delta_2} + \frac{R_2}{1 - \delta_2}. \quad (125)$$

From these the rate assertion of Theorem 4 follows.

APPENDIX C PROOF OF DISTRIBUTION INDEPENDENT SECURITY

We need to show that if Bob is honest, then (18) holds. In the proof we omit taking the maximum, but our argument is true for all joint distributions of (W_1, W_2) , hence the property follows.

We can almost identically follow the proof of Appendix VI-C. Similarly to (54) we have

$$I(W_1; Y_2^n S^n \Theta_2 | W_2) \leq I(W_1; Y_2^n S^n \Theta_2 U_C | W_2) = I(W_1; Y_2^n | Y_2^{n_1+n_2} S^n \Theta_2 U_C W_2). \quad (126)$$

The last step follows because given W_2 , variables Θ_A, Θ_2, S^n are independent of W_1 , further $Y_2^{n_1+n_2}, U^C$ are deterministic functions of $\Theta_A, \Theta_2, W_2, S^n$. The proofs of Lemmas 1 and 2 directly give us

$$I(K_B; Y_2^{n_1+n_2} S^{n_1+n_2}) \leq k_B e^{-c_3 \sqrt{k_1}}, \quad (127)$$

$$H(Y_2^n | Y_2^{n_1+n_2} S^n \Theta_2 U_C W_2) \leq \mathbb{E} \{M_B^C\}, \quad (128)$$

under the same conditions as defined in Lemmas 1 and 2, where $c_3 > 0$ is some constant. We still need to show that

$$H(Y_2^n | W_1 W_2 Y_2^{n_1+n_2} S^n \Theta_2 U_C) \geq \mathbb{E} \{ \min(k_B, M_B^C) \} - I(K_B; Y_2^{n_1+n_2} S^{n_1+n_2}) \quad (129)$$

holds. We can again follow the proof of Lemma 3, but we have to argue the step

$$I(K_B^C; W_1 W_2 Y_2^{n_1+n_2} S^n \Theta_2 U_C | 1_B^C) = I(K_B^C; Y_2^{n_1+n_2} S^{n_1+n_2} | 1_B^C), \quad (130)$$

where the independent and uniformly distributed property of W_2 was exploited when proving the lemma. To see that equation (130) is true under the modified protocol, consider that K_B^C is generated from a different set of random packets than K_B^C , so $K_B^C - Y_2^{n_1+n_2} - U_C$ is Markov-chain, and since (Θ_A, S^n) is generated independently of (W_1, W_2, Θ_2) , $K_B^C - (Y_2^{n_1+n_2}, 1_B^C, S^{n_1+n_2}) - (W_1, W_2, \Theta_2, U_C)$ has the Markov property too.

Having established the three key lemmas for the modified protocol, we can conclude the proof the same way as we have seen in Section VI. We omit the details to avoid repetitive arguments.

APPENDIX D PROOFS OF LEMMAS IN SECTION VII

We note that the order of proofs does not follow the order of appearance of the lemmas.

A. Proof of Lemma 7

Proof:

$$n(R_1 + R_2 + R_3) - \mathcal{E}_3(1 - \delta_1 \delta_2 \delta_3) \leq I(Y_1^n Y_2^n Y_3^n S^n; W_1 W_2 W_3) \quad (131)$$

$$= \sum_{i=1}^n I(Y_{1,i} Y_{2,i} Y_{3,i} S_i; W_1 W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}) \quad (132)$$

$$= \sum_{i=1}^n I(Y_{1,i} Y_{2,i} Y_{3,i}; W_1 W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1} S_i) \quad (133)$$

$$= \sum_{i=1}^n \Pr\{S_i \neq \emptyset\} I(Y_{1,i} Y_{2,i} Y_{3,i}; W_1 W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}, S_i \neq \emptyset) \quad (134)$$

$$= \sum_{i=1}^n I(X_i; W_1 W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}) (1 - \delta_1 \delta_2 \delta_3) \quad (135)$$

Here, the first inequality is Fano's inequality, besides, we exploited the independence property of S_i . This completes the proof of the first inequality of the lemma. Further, we also see that

$$I(Y_1^n Y_2^n Y_3^n S^n; W_1 W_2 W_3) \leq n(R_1 + R_2 + R_3), \quad (136)$$

which by a similar argument gives the second inequality of the lemma. ■

B. Proof of Lemma 5

Proof: From the same type of derivation as we apply in Lemma 7, we have that

$$\sum_{i=1}^n I(X_i; W_1 | Y_1^{i-1} S^{i-1}) \geq \frac{nR_1}{1 - \delta_1} - \mathcal{E}_1 \quad (137)$$

$$\sum_{i=1}^n I(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} S^{i-1}) \leq \frac{nR_1}{1 - \delta_1 \delta_2}. \quad (138)$$

Thus,

$$\frac{nR_1}{1 - \delta_1} - \mathcal{E}_1 \leq \sum_{i=1}^n I(X_i; W_1 | Y_1^{i-1} S^{i-1}) \quad (139)$$

$$= \sum_{i=1}^n I(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} S^{i-1}) + I(X_i; Y_2^{i-1} | Y_1^{i-1} S^{i-1}) - I(X_i; Y_2^{i-1} | Y_1^{i-1} S^{i-1} W_1) \quad (140)$$

$$\leq \sum_{i=1}^n I(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} S^{i-1}) + I(X_i; Y_2^{i-1} | Y_1^{i-1} S^{i-1}) \leq \frac{nR_1}{1 - \delta_1 \delta_2} + \sum_{i=1}^n I(X_i; Y_2^{i-1} | Y_1^{i-1} S^{i-1}) \quad (141)$$

■

C. Proof of Lemma 6

Proof: The proof follows the same kind of derivation as the proof of Lemma 5. We omit details to avoid repetition. ■

D. Proof of Lemma 8

Proof: We apply the shorthand $W^3 = W_1 W_2 W_3$.

$$0 \leq H(Y_1^n S^n | Y_2^n Y_3^n S^n W^3) = H(Y_1^{n-1} S^{n-1} | Y_2^n Y_3^n S^n W^3) + H(Y_{1,n} S_n | Y_1^{n-1} Y_2^n Y_3^n S^n W^3) \quad (142)$$

$$= H(Y_1^{n-1} S^{n-1} | Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) - I(Y_1^{n-1} S^{n-1}; Y_{2,n} Y_{3,n} S_n | Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) + H(Y_{1,n} | Y_1^{n-1} Y_2^n Y_3^n S^n W^3) \quad (143)$$

$$= H(Y_1^{n-1} S^{n-1} | Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) - I(Y_1^{n-1} S^{n-1}; Y_{2,n} Y_{3,n} | Y_2^{n-1} Y_3^{n-1} S^{n-1} S_n W^3) + H(Y_{1,n} | Y_1^{n-1} Y_2^n Y_3^n S^n W^3) \quad (144)$$

$$= H(Y_1^{n-1} S^{n-1} | Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) - I(Y_1^{n-1} S^{n-1}; Y_{2,n} Y_{3,n} | Y_2^{n-1} Y_3^{n-1} S^{n-1} S_n \notin \{\emptyset, \{1\}\} W^3) \Pr\{S_n \notin \{\emptyset, \{1\}\}\} \quad (145)$$

$$+ H(Y_{1,n} | Y_1^{n-1} Y_2^n Y_3^n S^n, S_n = \{1\}, W^3) \Pr\{S_n = \{1\}\} \quad (146)$$

$$= H(Y_1^{n-1} S^{n-1} | Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) - I(Y_1^{n-1} S^{n-1}; X_n | Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) (1 - \delta_2 \delta_3) \quad (147)$$

$$+ H(X_n | Y_1^{n-1} Y_2^{n-1} Y_3^{n-1} S^{n-1} W^3) (1 - \delta_1) \delta_2 \delta_3 \quad (148)$$

We do the same steps recursively to obtain the statement of the lemma. ■

E. Proof of Lemma 9

Proof: From (6), we have

$$\epsilon > I(Y_2^n Y_3^n S^n; W_1) = \sum_{i=1}^n I(X_i; W_1 | Y_2^{i-1} Y_3^{i-1} S^{i-1}) (1 - \delta_2 \delta_3) \quad (149)$$

We omitted the intermediate steps that are in the same vein as in the proof of Lemma 7. ■

F. Proof of Lemma 10

Proof:

$$\begin{aligned} \sum_{i=1}^n I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1 W_2 W_3) = \\ \sum_{i=1}^n I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) - I(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) + I(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) \end{aligned} \quad (150)$$

From (4) and Fano's inequality we have

$$I(Y_2^n Y_3^n S^n; W_2 W_3 | W_1) \leq I(Y_2^n Y_3^n S^n; W_2 W_3) - (h_2(\epsilon) + \epsilon). \quad (151)$$

We expand these terms the same way as we did in the proof of Lemma 7, and we can write

$$\sum_{i=1}^n I(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) \leq \sum_{i=1}^n I(X_i; W_2 W_3 | Y_2^{i-1} Y_3^{i-1} S^{i-1}) + \mathcal{E}_5. \quad (152)$$

From the independence property of the messages

$$\sum_{i=1}^n I(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) \geq \sum_{i=1}^n I(X_i; W_2 W_3 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}), \quad (153)$$

where $\mathcal{E}_5 = \frac{h_2(\epsilon) + \epsilon}{1 - \delta_2 \delta_3}$. This enables us to use the same idea as in Lemma 7 to bound these terms. Doing so gives us

$$\sum_{i=1}^n I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1 W_2 W_3) \geq \sum_{i=1}^n I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) - \frac{n(R_2 + R_3)}{1 - \delta_2 \delta_3} + \frac{n(R_2 + R_3)}{1 - \delta_1 \delta_2 \delta_3} - \mathcal{E}'_2 - \mathcal{E}_5, \quad (154)$$

where $\mathcal{E}'_2 = \frac{h_2(2\epsilon)+2\epsilon}{1-\delta_2\delta_3}$. It remains to give a bound on term (154). From Lemma 9 and after a few basic steps we can arrive to

$$\mathcal{E}_4 > \sum_{i=1}^n I(X_i; W_1 | Y_2^{i-1} Y_3^{i-1} S^{i-1}) \quad (155)$$

$$= \sum_{i=1}^n -I(X_i; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1} W_1) + I(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}) + I(X_1; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1}). \quad (156)$$

From a similar result as in Lemma 7:

$$I(X_i; W_1 | Y_1^{i-1} Y_2^{i-1} Y_3^{i-1} S^{i-1}) \geq \frac{nR_1}{1-\delta_1\delta_2\delta_3} - \mathcal{E}_6, \quad (157)$$

where $\mathcal{E}_6 = \frac{h_2(\epsilon)+\epsilon}{1-\delta_1\delta_2\delta_3}$. Further, a symmetric result to Lemma 6 shows:

$$I(X_1; Y_1^{i-1} | Y_2^{i-1} Y_3^{i-1} S^{i-1}) \frac{n(R_2 + R_3)}{1-\delta_2\delta_3} - \frac{n(R_2 + R_3)}{1-\delta_1\delta_2\delta_3} - \mathcal{E}'_2. \quad (158)$$

Applying these bounds in (156) and then substituting back to (154) results the claim of the lemma. \blacksquare

APPENDIX E PROOF OF LEMMA 11

Proof: As a first step we define

$$\mathbf{Adv}_{\text{dis}}^{*ss} = \max_{f, P_{W_1}, w_2, \sigma} \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \Theta_2, \sigma, w_2) = f(W_1, w_2) \} - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, w_2) = f(W_1, w_2) \} \right\}. \quad (159)$$

As opposed to \mathbf{Adv}^{ss} here w_2 is not a random variable but a constant value from \mathcal{W}_2 . Clearly,

$$\mathbf{Adv}_{\text{dis}}^{*ss} \leq \mathbf{Adv}_{\text{dis}}^{ss}, \quad (160)$$

because W_2 taking the value w_2 with probability 1 is a particular joint distribution W_1, W_2 can take, so the scope of the maximization is restricted. We show that $\mathbf{Adv}_{\text{dis}}^{*ss} = \mathbf{Adv}_{\text{dis}}^{ss}$.

$$\mathbf{Adv}_{\text{dis}}^{ss} = \max_{f, P_{W_1}, w_2, \sigma} \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \Theta_2, \sigma, W_2) = f(W_1, W_2) \} - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, W_2) = f(W_1, W_2) \} \right\} \quad (161)$$

$$= \max_{f, P_{W_1}, w_2, \sigma} \sum_{w_2} p_{W_2}(w_2) \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \Theta_2, \sigma, W_2) = f(W_1, W_2) | W_2 = w_2 \} \right. \quad (162)$$

$$\left. - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, \sigma, W_2) = f(W_1, W_2) | W_2 = w_2 \} \right\} \quad (163)$$

$$= \max_{w_2^*, f, P_{W_1|W_2=w_2^*}, \sigma} \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \Theta_2, \sigma, w_2^*) = f(W_1, w_2^*) \} - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, w_2^*) = f(W_1, w_2^*) \} \right\} \quad (164)$$

$$= \max_{f, P_{W_1}, w_2^*, \sigma} \left\{ \max_{\mathcal{A}} \Pr \{ \mathcal{A}(Y_2^n, S^n, \Theta_2, \sigma, w_2^*) = f(W_1, w_2^*) \} - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, w_2^*) = f(W_1, w_2^*) \} \right\} \quad (165)$$

$$= \mathbf{Adv}_{\text{dis}}^{*ss}. \quad (166)$$

where the second step follows because there is a certain value w_2^* of W_2 that maximizes the expression inside $\{\dots\}$, and moreover this expression depends on $P_{W_1|W_2}$ only through $P_{W_1|W_2=w_2}$, hence a maximizing joint distribution of W_1, W_2 is when W_2 takes this particular value with probability 1.

We continue the proof in two steps, first we define a notion of distinguishing security applicable for jointly distributed messages by extending a similar definition in [13] and show its equivalence with the above definition of semantic security. Then we show equivalence between this notion of distinguishing security and distribution independent security as defined by $\mathbf{Adv}_{\text{dis}}^{\text{mis}}$.

We define a notion corresponding to distinguishing security by defining the adversarial advantage:

$$\mathbf{Adv}_{\text{dis}}^{\text{ds}} = \max_{\mathcal{A}, w_1^0, w_1^1, w_2, \sigma} 2 \Pr \{ \mathcal{A}(w_1^0, w_1^1, w_2, {}^b Y_2^n, S^n, \Theta_2, \sigma) = b \} - 1, \quad (167)$$

where $w_1^0, w_1^1 \in \mathcal{W}_1$ are possible messages, similarly $w_2 \in \mathcal{W}_2$, b is a variable uniformly distributed over $\{0, 1\}$ and is independent of all other variables, while ${}^b Y_2^n$ is Calvin's observation given $W_1 = w_1^b$.

Distinguishing security defined by $\mathbf{Adv}_{\text{dis}}^{\text{ds}}$ is equivalent to semantic security as defined by $\mathbf{Adv}_{\text{dis}}^{*ss}$ and hence equivalently as defined by $\mathbf{Adv}_{\text{dis}}^{ss}$. To show that distinguishing security implies semantic security, we can almost identically follow the proof of Theorem 5 from [13], with a slight difference that a conditioning on W_2 appears. Given an adversary \mathcal{A}_{ss} attacking semantic security, we construct an adversary \mathcal{A}_{ds} attacking distinguishing security as follows: \mathcal{A}_{ds} outputs 1, if the adversary

attacking semantic security \mathcal{A}_{ss} gives as output $f(w_1^1, w_2)$, otherwise it returns 0. Then, if W_1^0 and W_1^1 are i.i.d. both having the same distribution as W_1 , then

$$\Pr \{ \mathcal{A}_{ds}(W_1^0, W_1^1, W_2, {}^1Y_2^n, S^n, \Theta_2, \sigma) = 1 | W_2 = w_2 \} = \Pr \{ A_{ss}(Y_2^n, S^n, \Theta_2, \sigma, W_2) = f(W_1, W_2) | W_2 = w_2 \} \quad (168)$$

$$\Pr \{ \mathcal{A}_{ds}(W_1^0, W_1^1, W_2, {}^0Y_2^n, S^n, \Theta_2, \sigma) = 1 | W_2 = w_2 \} \leq \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, W_2) = f(W_1, W_2) | W_2 = w_2 \}. \quad (169)$$

Finishing the derivation as in [13] we get

$$\begin{aligned} \Pr \{ A_{ss}(Y_2^n, S^n, \Theta_2, \sigma, W_2) = f(W_1, W_2) | W_2 = w_2 \} - \max_S \Pr \{ \mathcal{S}(P_{W_1}, f, W_2) = f(W_1, W_2) | W_2 = w_2 \} \\ \leq \max_{w_1^0, w_1^1, w_2, \mathcal{A}_{ds}, \sigma} 2 \Pr \{ \mathcal{A}_{ds}(w_1^0, w_1^1, w_2, {}^bY_2^n, S^n, \Theta_2, \sigma) = b \} - 1 \end{aligned}$$

for all $P_{W_1}, f, \mathcal{A}_{ss}, \sigma$, hence taking the maximum over these variables on the LHS and over w_2 on both sides keeps the inequality. This establishes that

$$\mathbf{Adv}_{\text{dis}}^{\text{ss}} = \mathbf{Adv}_{\text{dis}}^{*\text{ss}} \leq \mathbf{Adv}_{\text{dis}}^{\text{ds}} \leq 2\mathbf{Adv}_{\text{dis}}^{\text{ss}}. \quad (170)$$

The other direction of implication is a straightforward consequence of the definitions, the scope of maximization in $\mathbf{Adv}_{\text{dis}}^{\text{ds}}$ is a subset of that of $\mathbf{Adv}_{\text{dis}}^{\text{ss}}$, in case of $\mathbf{Adv}_{\text{dis}}^{\text{ds}}$ f is a function that computes b , while P_{W_1, W_2} is such that W_1 uniformly takes the two values w_1^0 and w_1^1 and independently W_2 takes w_2 with probability 1.

What remains to show is that distinguishing security defined by $\mathbf{Adv}_{\text{dis}}^{\text{ds}}$ is equivalent to distribution independent security as defined by $\mathbf{Adv}_{\text{dis}}^{\text{mis}}$. Clearly, for any particular value of w_2 ,

$$\mathbf{Adv}_{\text{dis}}^{\text{mis}} \geq \max_{P_{W_1}, \sigma} I(W_1; Y_2^n S^n \Theta_2 | W_2 = w_2). \quad (171)$$

If we fix w_2 for the scheme, we can directly invoke Theorem 5 from [13] which proves that

$$\max_{\mathcal{A}, w_1^0, w_1^1, \sigma} 2 \Pr \{ \mathcal{A}(w_1^0, w_1^1, w_2, {}^bY_2^n, S^n, \Theta_2, \sigma) = b \} - 1 \leq \sqrt{2 \max_{P_{W_1}, \sigma} I(W_1; Y_2^n S^n \Theta_2 | W_2 = w_2)} \leq \sqrt{2 \cdot \mathbf{Adv}_{\text{dis}}^{\text{mis}}}, \quad (172)$$

which holds for every w_2 , so we can take the maximum in w_2 on the LHS, which gives in turn

$$\mathbf{Adv}_{\text{dis}}^{\text{ds}} \leq \sqrt{2 \cdot \mathbf{Adv}_{\text{dis}}^{\text{mis}}} \quad (173)$$

showing that the distribution independent security implies distinguishing security. The other direction is also true. We can apply the same type of argument as when showing $\mathbf{Adv}_{\text{dis}}^{\text{ss}} = \mathbf{Adv}_{\text{dis}}^{*\text{ss}}$ to get:

$$\mathbf{Adv}_{\text{dis}}^{\text{mis}} = \max_{P_{W_1, W_2}, \sigma} I(W_1; Y_2^n S^n \Theta_2 | W_2) \quad (174)$$

$$= \max_{w_2, P_{W_1}, \sigma} I(W_1; Y_2^n S^n \Theta_2 | W_2 = w_2). \quad (175)$$

Let us denote

$$\mathbf{Adv}^{\text{ds}}(w_2) = \max_{\mathcal{A}, w_1^0, w_1^1, \sigma} 2 \Pr \{ \mathcal{A}(w_1^0, w_1^1, w_2, {}^bY_2^n, S^n, \Theta_2, \sigma) = b \} - 1, \quad (176)$$

We can apply Theorem 4.9 from [13] with a conditioning on $W_2 = w_2$, which implies that for any w_2 :

$$\max_{P_{W_1}, \sigma} I(W_1; Y_2^n S^n \Theta_2 | W_2 = w_2) \leq 2\mathbf{Adv}^{\text{ds}}(w_2) \log \left(\frac{2^n}{\mathbf{Adv}^{\text{ds}}(w_2)} \right). \quad (177)$$

Since the above is true for any w_2 , we can take the maximum in w_2 on both sides resulting

$$\mathbf{Adv}_{\text{dis}}^{\text{mis}} \leq 2\mathbf{Adv}^{\text{ds}} \log \left(\frac{2^n}{\mathbf{Adv}^{\text{ds}}} \right). \quad (178)$$

This completes the proof that distribution independent security is equivalent to semantic security defined by $\mathbf{Adv}_{\text{dis}}^{\text{ss}}$. ■